

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/010913

International filing date: 31 March 2005 (31.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/559,202
Filing date: 01 April 2004 (01.04.2004)

Date of receipt at the International Bureau: 12 August 2005 (12.08.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1352381

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 02, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/559,202

FILING DATE: *April 01, 2004*

RELATED PCT APPLICATION NUMBER: *PCT/US05/10913*



Certified by

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

16085 U.S. PTO
040104

PTO/SB/16 (08-03)

Approved for use through 7/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EL 961008121 US

22154 U.S. PTO
607559202

040104

INVENTOR(S)					
Given Name (first and middle [if any])		Family Name or Surname		Residence (City and either State or Foreign Country)	
Steven Andreas Stefan M.		MAH BRAUN KAMMERER		San Diego, California San Diego, California San Diego, California	
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto					
TITLE OF THE INVENTION (500 characters max)					
METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF					
Direct all correspondence to: CORRESPONDENCE ADDRESS					
<input checked="" type="checkbox"/> Customer Number:		25225			
OR					
<input type="checkbox"/> Firm or Individual Name					
Address					
City		State		Zip	
Country		Telephone		Fax	
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/>	Specification Number of Pages	240	<input type="checkbox"/>	CD(s), Number	
<input checked="" type="checkbox"/>	Drawing(s) Number of Sheets	2	<input checked="" type="checkbox"/>	Other	Return Receipt Postcard
<input checked="" type="checkbox"/>	Application Data Sheet. See 37 CFR 1.76 (4 pages) (specify):				
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT					
<input checked="" type="checkbox"/>	Applicant claims small entity status. See 37 CFR 1.27.				FILING FEE AMOUNT (\$) 80.00
<input type="checkbox"/>	A check or money order is enclosed to cover the filing fees.				
<input checked="" type="checkbox"/>	The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 03-1952				
<input type="checkbox"/>	Payment by credit card. Form PTO-2038 is attached.				
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input checked="" type="checkbox"/>	No	<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are:			

[Page 1 of 2]

Respectfully submitted,

Date April 1, 2004

SIGNATURE

TYPED OR
PRINTED NAME

Bruce D. Grant

REGISTRATION NO.
(if appropriate)

47,608

TELEPHONE

(858) 720-7962

Docket Number:

524593008700

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as Express Mail, Airbill No. EL 961008121 US, in an envelope addressed to: Mail Stop Provisional Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on the date shown below.

Dated:

4/1/04

Signature:

(Deborah Wykes)

SD-189429

PROVISIONAL APPLICATION COVER SHEET
Additional Page

PTO/SB/16 (08-03)

Approved for use through 07/31/06. OMB 0651-0032

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Docket Number 524593008700

INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle [if any])	Family or Surname	Residence (City and either State or Foreign Country)
Matthew Roberts	NELSON	San Marcos, California
Rikard Henry	RENELAND	San Diego, California
Maria L.	LANGDOWN	San Diego, California

[Page 2 of 2]

METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF

Field of the Invention

[0001] The invention relates to genetic methods for identifying risk of osteoarthritis and treatments that specifically target such diseases.

Background

[0002] Osteoarthritis (OA) is a chronic disease usually affecting weight-bearing synovial joints. There are approximately 20 million Americans affected by OA and it is the leading cause of disability in the United States. In addition to extensive human suffering, OA also accounts for nearly all knee replacements and more than half of all hip replacements in the United States. Despite its prevalence, OA is poorly understood and there are few treatments available besides anti-inflammatory drugs and joint replacement.

[0003] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA can be as disabling as any cardiovascular disease except stroke.

[0004] OA is characterized by the breakdown of cartilage in joints. Cartilage in joints cushions the ends of bones, and cartilage breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. *Arthritis Rheum.* 32:241-246 (1989)).

Summary

[0005] It has been discovered that certain polymorphic variations in human genomic DNA are associated with osteoarthritis. In particular, polymorphic variants in loci containing *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* and *ERG* regions in human genomic DNA have been associated with risk of osteoarthritis. The *PSMB1*, *TBP* and *PDCD2* regions are located in a larger region referred to herein as the *chrom 6* region.

[0006] Thus, featured herein are methods for identifying a subject at risk of osteoarthritis and/or a risk of osteoarthritis in a subject, which comprise detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in or around the loci described herein in a human nucleic acid sample. In an embodiment, two or more polymorphic variations are detected in two or more regions, of which one or more is a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* region. In certain

embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected.

[0007] Also featured are nucleic acids that include one or more polymorphic variations associated with occurrence of osteoarthritis, as well as polypeptides encoded by these nucleic acids. In addition, provided are methods for identifying candidate therapeutic molecules for treating osteoarthritis, as well as methods for treating osteoarthritis in a subject by identifying a subject at risk of osteoarthritis and treating the subject with a suitable prophylactic, treatment or therapeutic molecule.

[0008] Also provided are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid, with a RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid designed from a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence. In an embodiment, the RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid is designed from a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence that includes one or more polymorphic variations associated with osteoarthritis, and in some instances, specifically interacts with such a nucleotide sequence. Further, provided are arrays of nucleic acids bound to a solid surface, in which one or more nucleic acid molecules of the array have a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence, or a fragment or substantially identical nucleic acid thereof, or a complementary nucleic acid of the foregoing. Featured also are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* polypeptide, with an antibody that specifically binds to the polypeptide. In an embodiment, the antibody specifically binds to an epitope in the polypeptide that includes a non-synonymous amino acid modification associated with osteoarthritis (e.g., results in an amino acid substitution in the encoded polypeptide associated with osteoarthritis). In certain embodiments, the antibody selectively binds to an epitope in the *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* polypeptide having an amino acid associated with osteoarthritis.

Brief Description of the Drawings

[0009] Figures 1A-1D show proximal SNPs in *chrom 6*, *ELP3*, *CHDC1* and *ERG* regions of genomic DNA, respectively. The position of each SNP in the chromosome is shown on the x-axis and the y-axis provides the negative logarithm of the p-value comparing the estimated allele to that of the control group. Also shown in the figures are exons and introns of the regions in the approximate chromosomal positions.

Detailed Description

[0010] It has been discovered that a polymorphic variant in a locus containing a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* region is associated with occurrence of osteoarthritis in subjects. Thus,

detecting genetic determinants associated with an increased risk of osteoarthritis occurrence can lead to early identification of a predisposition to osteoarthritis and early prescription of preventative measures. Also, associating a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* polymorphic variant with osteoarthritis has provided new targets for screening molecules useful in treatments of osteoarthritis.

Osteoarthritis and Sample Selection

[0011] Osteoarthritis (OA), or degenerative joint disease, is one of the oldest and most common types of arthritis. It is characterized by the breakdown of the joint's cartilage. Cartilage is the part of the joint that cushions the ends of bones, and its breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. Arthritis Rheum. 32:241-246 (1989)).

[0012] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA, can be as disabling as any cardiovascular disease except stroke.

[0013] Osteoarthritis affects an estimated 20.7 million Americans, mostly after age 45, with women more commonly affected than men. Physicians make a diagnosis of OA based on a physical exam and history of symptoms. X-rays are used to confirm diagnosis. Most people over 60 reflect the disease on X-ray, and about one-third have actual symptoms.

[0014] There are many factors that can cause OA. Obesity may lead to osteoarthritis of the knees. In addition, people with joint injuries due to sports, work-related activity or accidents may be at increased risk of developing OA.

[0015] Genetics has a role in the development of OA. Some people may be born with defective cartilage or with slight defects in the way that joints fit together. As a person ages, these defects may cause early cartilage breakdown in the joint or the inability to repair damaged or deteriorated cartilage in the joint.

[0016] Inclusion or exclusion of samples for an osteoarthritis pool may be based upon the following criteria: ethnicity (e.g., samples derived from an individual characterized as Caucasian); parental ethnicity (e.g., samples derived from an individual of British paternal and maternal descent); relevant phenotype information for the individual (e.g., case samples derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic). Control samples may be selected based on relevant phenotype information for the individual (e.g., derived from individuals free of OA at several sites (knee, hand, hip etc)); and no family history of OA and/or rheumatoid arthritis. Additional phenotype information collected for both cases and controls may include

age of the individual, gender, family history of OA, diagnosis with osteoarthritis (joint location of OA, date of primary diagnosis, age of individual as of primary diagnosis), knee history (current symptoms, any major knee injury, meniscectomy, knee replacement surgery, age of surgery), HRT history, osteoporosis diagnosis.

[0017] Based in part upon selection criteria set forth above, individuals having osteoarthritis can be selected for genetic studies. Also, individuals having no history of osteoarthritis often are selected for genetic studies, as described hereafter.

Polymorphic Variants Associated with Osteoarthritis

[0018] A genetic analysis provided herein linked osteoarthritis with polymorphic variant nucleic acid sequences in the human genome. As used herein, the term “polymorphic site” refers to a region in a nucleic acid at which two or more alternative nucleotide sequences are observed in a significant number of nucleic acid samples from a population of individuals. A polymorphic site may be a nucleotide sequence of two or more nucleotides, an inserted nucleotide or nucleotide sequence, a deleted nucleotide or nucleotide sequence, or a microsatellite, for example. A polymorphic site that is two or more nucleotides in length may be 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more, 20 or more, 30 or more, 50 or more, 75 or more, 100 or more, 500 or more, or about 1000 nucleotides in length, where all or some of the nucleotide sequences differ within the region. A polymorphic site is often one nucleotide in length, which is referred to herein as a “single nucleotide polymorphism” or a “SNP.”

[0019] Where there are two, three, or four alternative nucleotide sequences at a polymorphic site, each nucleotide sequence is referred to as a “polymorphic variant” or “nucleic acid variant.” Where two polymorphic variants exist, for example, the polymorphic variant represented in a minority of samples from a population is sometimes referred to as a “minor allele” and the polymorphic variant that is more prevalently represented is sometimes referred to as a “major allele.” Many organisms possess a copy of each chromosome (*e.g.*, humans), and those individuals who possess two major alleles or two minor alleles are often referred to as being “homozygous” with respect to the polymorphism, and those individuals who possess one major allele and one minor allele are normally referred to as being “heterozygous” with respect to the polymorphism. Individuals who are homozygous with respect to one allele are sometimes predisposed to a different phenotype as compared to individuals who are heterozygous or homozygous with respect to another allele.

[0020] In genetic analysis that associate polymorphic variants with osteoarthritis, samples from individuals having osteoarthritis and individuals not having osteoarthritis often are allelotyped and/or genotyped. The term “allelotype” as used herein refers to a process for determining the allele frequency for a polymorphic variant in pooled DNA samples from cases and controls. By pooling DNA from each group, an allele frequency for each SNP in each group is calculated. These allele frequencies are then

compared to one another. The term “genotyped” as used herein refers to a process for determining a genotype of one or more individuals, where a “genotype” is a representation of one or more polymorphic variants in a population.

[0021] A genotype or polymorphic variant may be expressed in terms of a “haplotype,” which as used herein refers to two or more polymorphic variants occurring within genomic DNA in a group of individuals within a population. For example, two SNPs may exist within a gene where each SNP position includes a cytosine variation and an adenine variation. Certain individuals in a population may carry one allele (heterozygous) or two alleles (homozygous) having the gene with a cytosine at each SNP position. As the two cytosines corresponding to each SNP in the gene travel together on one or both alleles in these individuals, the individuals can be characterized as having a cytosine/cytosine haplotype with respect to the two SNPs in the gene.

[0022] As used herein, the term “phenotype” refers to a trait which can be compared between individuals, such as presence or absence of a condition, a visually observable difference in appearance between individuals, metabolic variations, physiological variations, variations in the function of biological molecules, and the like. An example of a phenotype is occurrence of osteoarthritis.

[0023] Researchers sometimes report a polymorphic variant in a database without determining whether the variant is represented in a significant fraction of a population. Because a subset of these reported polymorphic variants are not represented in a statistically significant portion of the population, some of them are sequencing errors and/or not biologically relevant. Thus, it is often not known whether a reported polymorphic variant is statistically significant or biologically relevant until the presence of the variant is detected in a population of individuals and the frequency of the variant is determined. Methods for detecting a polymorphic variant in a population are described herein, specifically in Example 2. A polymorphic variant is statistically significant and often biologically relevant if it is represented in 5% or more of a population, sometimes 10% or more, 15% or more, or 20% or more of a population, and often 25% or more, 30% or more, 35% or more, 40% or more, 45% or more, or 50% or more of a population.

[0024] A polymorphic variant may be detected on either or both strands of a double-stranded nucleic acid. Also, a polymorphic variant may be located within an intron or exon of a gene or within a portion of a regulatory region such as a promoter, a 5′ untranslated region (UTR), a 3′ UTR, and in DNA (*e.g.*, genomic DNA (gDNA) and complementary DNA (cDNA)), RNA (*e.g.*, mRNA, tRNA, and rRNA), or a polypeptide. Polymorphic variations may or may not result in detectable differences in gene expression, polypeptide structure, or polypeptide function.

[0025] It was determined that polymorphic variations associated with an increased risk of osteoarthritis existed in the *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* regions. In certain embodiments, polymorphic variants at positions rs756519, rs1042327, rs8770, rs1563055, rs912428 and rs1888475 in the human genome were associated with an increased risk of osteoarthritis, and in specific

embodiments, a thymine at position rs756519, a cytosine at position rs1042327, a cytosine at position rs8770, a thymine at position rs1563055, a thymine at position rs912428 and an adenine at position rs1888475 were associated with an increased risk of osteoarthritis.

[0026] Polymorphic variants in and around the *chrom 6* region were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 1 selected from the group consisting of 229, 6310, 11840, 11870, 12064, 13392, 16354, 16559, 16935, 17616, 17737, 18321, 18453, 18811, 20020, 21662, 23197, 23446, 24339, 25504, 27174, 28008, 29294, 29759, 30832, 44512, 44850, 45884, 46345, 48589, 53371, 53911, 53990, 55152, 55667, 58952, 59315, 60029, 61477, 62988, 63090, 64021, 65685, 70220, 70323, 70959, 73436, 82945, 82958, 82961, 82964, 82965, 83006, 83025, 83034, 83074, 83132, 83155, 83172, 83174, 83206, 83216, 83234, 83252, 83260, 83263, 83296, 83319, 83322, 83324, 83357, 83375, 83381, 83389, 83443, 83499, 83545, 83566, 83591, 83619, 83698, 83780, 83784, 83826, 83832, 83852, 86297, 86315, 86420, 86460, 86714, 86718, 86736, 86753, 86766, 88162, 88218, 88246, 88255, 88309, 88310, 88471, 88619, 88904, 89044, 90531, 90534, 90613 and 46252. Polymorphic variants at the following positions in SEQ ID NO: 1 in particular were associated with an increased risk of osteoarthritis: 229, 6310, 16559, 18453, 25504, 27174, 30832, 44850, 45884, 48589, 61477, 82961 and 46252, with specific cmbodiments directed to variants at positions 229, 16559, 44850 and/or 46252. In particular, the following polymorphic variants in SEQ ID NO: 1 were associated with risk of osteoarthritis: a thymine at position 229, a guanine at position 6310, a thymine at position 16559, an adenine at position 18453, an adenine at position 25504, an adenine at position 27174, an adenine at position 30832, a guanine at position 44850, an adenine at position 45884, an adenine at position 48589, a cytosine at position 61477, a cytosine at position 82961 and a thymine at position 46252.

[0027] Polymorphic variants in and around the *ELP3* region were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 2 selected from the group consisting of 211, 473, 1536, 5639, 17186, 17335, 25029, 25111, 28811, 28863, 30809, 40985, 45147, 45282, 46168, 46328, 49077, 51925, 52141, 52168, 60852, 62468, 65572, 79089, 79541, 79790, 90843, 90978, 91052, 91131, 91132, 94439 and 94621. Polymorphic variants at the following positions in SEQ ID NO: 2 in particular were associated with an increased risk of osteoarthritis: 40985, 46168, 51925 and 52168. In particular, the following polymorphic variants in SEQ ID NO: 2 were associated with risk of osteoarthritis: a cytosine at position 40985, a guanine at position 46168, a thymine at position 51925 and a cytosine at position 52168.

[0028] Polymorphic variants in and around the *CHDC1* region were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 3 selected from the group consisting of 243, 10208, 15049, 15111, 15272, 15287, 15326, 15327, 17038, 19391, 21702, 22431, 22881, 27744, 32564, 32698, 33104, 33181, 33256, 33543, 35567, 40085, 40482, 45641, 46059, 48504, 48919, 49693, 49874, 50020, 50616, 50719, 55511, 65533, 70529, 75591, 77266, 80368, 82475, 92462,

92480, 95819 and 96275. Polymorphic variants at the following positions in SEQ ID NO: 3 in particular were associated with an increased risk of osteoarthritis: 15111, 45641, 46059, 49693, 49874, 50020, 50719, 70529, 82475, 92462, 92480 and 96275, with specific embodiments directed to variants at positions 82475 and/or 92462. In particular, the following polymorphic variants in SEQ ID NO: 3 were associated with risk of osteoarthritis: a guanine at position 15111, a thymine at position 45641, an adenine at position 46059, a cytosine at position 49693, an adenine at position 49874, an adenine at position 50020, a guanine at position 50719, an adenine at position 70529, an adenine at position 82475, a thymine at position 92462, a thymine at position 92480 and a cytosine at position 96275.

[0029] Polymorphic variants in and around the *ERG* region were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 4 selected from the group consisting of 231, 882, 960, 1194, 1530, 1673, 2096, 2285, 5873, 7256, 7988, 8222, 8381, 8814, 8915, 9642, 9902, 10619, 10927, 11032, 14377, 15608, 15928, 16296, 17598, 19272, 20084, 20577, 28051, 29466, 29530, 29987, 30012, 30322, 32216, 32516, 32544, 32746, 33137, 33538, 33798, 33802, 33964, 34132, 34210, 34317, 34499, 34753, 34845, 35335, 36423, 36450, 36481, 38447, 38784, 39387, 39458, 39822, 40305, 40869, 40926, 41010, 41134, 41984, 42172, 42753, 43011, 43176, 43320, 43381, 44142, 44383, 44726, 45087, 45141, 45359, 45421, 45456, 45467, 45486, 45709, 45716, 47626, 49413, 49796, 49962, 50075, 50093, 50571, 50615, 50780, 50851, 51459, 53193, 53702, 53736, 53795, 54109, 54126, 54230, 54894, 55455, 55499, 56522, 56662, 56954, 57267, 58282, 58916, 59544, 59666, 59913, 66846, 67245, 67652, 67955, 67966, 68420, 70226, 70810, 72246, 73330, 73457, 74389, 74638, 74640, 75358, 75952, 76098, 77836, 78449, 78507, 80031, 81695, 82775, 82795, 84611, 84657, 84693, 85020, 85048, 85100, 85325, 85452, 85868, 85936, 85990, 86139, 86497, 87236, 87248, 87533, 87912, 88108, 88494, 89598, 90235, 91287, 91359, 92384, 92410, 92900, 94495, 94512, 97777 and 98333. Polymorphic variants at the following positions in SEQ ID NO: 4 in particular were associated with an increased risk of osteoarthritis: 1673, 20577, 33137, 39822, 45716, 49962, 51459, 54894, 55455, 55499, 58282, 68420 and 80031, with specific embodiments directed to variants at positions 33137, 55499 and/or 58282. In particular, the following polymorphic variants in SEQ ID NO: 4 were associated with risk of osteoarthritis: a guanine at position 1673, a thymine at position 20577, a guanine at position 33137, a guanine at position 39822, an adenine at position 45716, a guanine at position 49962, an adenine at position 51459, a cytosine at position 54894, an adenine at position 55455, an adenine at position 55499, a guanine at position 58282, an adenine at position 68420 and a thymine at position 80031.

[0030] Based in part upon analyses summarized in Figures 1A-1D, regions with significant association have been identified in loci associated with osteoarthritis. Any polymorphic variants associated with osteoarthritis in a region of significant association can be utilized for embodiments described herein. For example, polymorphic variants in a region spanning chromosome positions 170719500 to 170766500 (approximately 47,000 nucleotides in length) in a *chrom 6* region, spanning

chromosome positions 27963000 to 27983000 (approximately 20,000 nucleotides in length) in a *ELP3* region, spanning chromosome positions 44962000 to 45013000 (approximately 51,000 nucleotides in length) in a *CHDC1* region, and spanning chromosome positions 38830000 to 38844000 (approximately 14,000 nucleotides in length) in a *ERG* region, have significant association (chromosome positions are within NCBI's Genome build 34).

Additional Polymorphic Variants Associated with Osteoarthritis

[0031] Also provided is a method for identifying polymorphic variants proximal to an incident, founder polymorphic variant associated with osteoarthritis. Thus, featured herein are methods for identifying a polymorphic variation associated with osteoarthritis that is proximal to an incident polymorphic variation associated with osteoarthritis, which comprises identifying a polymorphic variant proximal to the incident polymorphic variant associated with osteoarthritis, where the incident polymorphic variant is in a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence. The nucleotide sequence often comprises a polynucleotide sequence selected from the group consisting of (a) a polynucleotide sequence of SEQ ID NO: 1-12; (b) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence encoded by a polynucleotide sequence of SEQ ID NO: 1-12; and (c) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1 or a polynucleotide sequence 90% or more identical to the polynucleotide sequence of SEQ ID NO: 1-12. The presence or absence of an association of the proximal polymorphic variant with osteoarthritis then is determined using a known association method, such as a method described in the Examples hereafter. In an embodiment, the incident polymorphic variant is a polymorphic variant associated with osteoarthritis described herein. In another embodiment, the proximal polymorphic variant identified sometimes is a publicly disclosed polymorphic variant, which for example, sometimes is published in a publicly available database. In other embodiments, the polymorphic variant identified is not publicly disclosed and is discovered using a known method, including, but not limited to, sequencing a region surrounding the incident polymorphic variant in a group of nucleic samples. Thus, multiple polymorphic variants proximal to an incident polymorphic variant are associated with osteoarthritis using this method.

[0032] The proximal polymorphic variant often is identified in a region surrounding the incident polymorphic variant. In certain embodiments, this surrounding region is about 50 kb flanking the first polymorphic variant (e.g. about 50 kb 5' of the first polymorphic variant and about 50 kb 3' of the first polymorphic variant), and the region sometimes is composed of shorter flanking sequences, such as flanking sequences of about 40 kb, about 30 kb, about 25 kb, about 20 kb, about 15 kb, about 10 kb, about 7 kb, about 5 kb, or about 2 kb 5' and 3' of the incident polymorphic variant. In other embodiments, the region is composed of longer flanking sequences, such as flanking sequences of about

55 kb, about 60 kb, about 65 kb, about 70 kb, about 75 kb, about 80 kb, about 85 kb, about 90 kb, about 95 kb, or about 100 kb 5' and 3' of the incident polymorphic variant.

[0033] In certain embodiments, polymorphic variants associated with osteoarthritis are identified iteratively. For example, a first proximal polymorphic variant is associated with osteoarthritis using the methods described above and then another polymorphic variant proximal to the first proximal polymorphic variant is identified (*e.g.*, publicly disclosed or discovered) and the presence or absence of an association of one or more other polymorphic variants proximal to the first proximal polymorphic variant with osteoarthritis is determined.

[0034] The methods described herein are useful for identifying or discovering additional polymorphic variants that may be used to further characterize a gene, region or loci associated with a condition, a disease (*e.g.*, osteoarthritis), or a disorder. For example, allelotyping or genotyping data from the additional polymorphic variants may be used to identify a functional mutation or a region of linkage disequilibrium. In certain embodiments, polymorphic variants identified or discovered within a region comprising the first polymorphic variant associated with osteoarthritis are genotyped using the genetic methods and sample selection techniques described herein, and it can be determined whether those polymorphic variants are in linkage disequilibrium with the first polymorphic variant. The size of the region in linkage disequilibrium with the first polymorphic variant also can be assessed using these genotyping methods. Thus, provided herein are methods for determining whether a polymorphic variant is in linkage disequilibrium with a first polymorphic variant associated with osteoarthritis, and such information can be used in prognosis/diagnosis methods described herein.

Isolated Nucleic Acids

[0035] Featured herein are isolated *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid variants depicted in SEQ ID NO: 1-12, and substantially identical nucleic acids thereof. A nucleic acid variant may be represented on one or both strands in a double-stranded nucleic acid or on one chromosomal complement (heterozygous) or both chromosomal complements (homozygous).

[0036] As used herein, the term "nucleic acid" includes DNA molecules (*e.g.*, a complementary DNA (cDNA) and genomic DNA (gDNA)) and RNA molecules (*e.g.*, mRNA, rRNA, siRNA and tRNA) and analogs of DNA or RNA, for example, by use of nucleotide analogs. The nucleic acid molecule can be single-stranded and it is often double-stranded. The term "isolated or purified nucleic acid" refers to nucleic acids that are separated from other nucleic acids present in the natural source of the nucleic acid. For example, with regard to genomic DNA, the term "isolated" includes nucleic acids which are separated from the chromosome with which the genomic DNA is naturally associated. An "isolated" nucleic acid is often free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and/or 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic

acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of 5' and/or 3' nucleotide sequences which flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. As used herein, the term "gene" refers to a nucleotide sequence that encodes a polypeptide.

[0037] Also included herein are nucleic acid fragments. These fragments often have a nucleotide sequence identical to a nucleotide sequence of SEQ ID NO: 1-12, a nucleotide sequence substantially identical to a nucleotide sequence of SEQ ID NO: 1-12, or a nucleotide sequence that is complementary to the foregoing. The nucleic acid fragment may be identical, substantially identical or homologous to a nucleotide sequence in an exon or an intron in a nucleotide sequence of SEQ ID NO: 1-12, and may encode a domain or part of a domain of a polypeptide. Sometimes, the fragment will comprises one or more of the polymorphic variations described herein as being associated with osteoarthritis. The nucleic acid fragment is often 50, 100, or 200 or fewer base pairs in length, and is sometimes about 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 2000, 3000, 4000, 5000, 10000, 15000, or 20000 base pairs in length. A nucleic acid fragment that is complementary to a nucleotide sequence identical or substantially identical to a nucleotide sequence in SEQ ID NO: 1-12 and hybridizes to such a nucleotide sequence under stringent conditions is often referred to as a "probe." Nucleic acid fragments often include one or more polymorphic sites, or sometimes have an end that is adjacent to a polymorphic site as described hereafter.

[0038] An example of a nucleic acid fragment is an oligonucleotide. As used herein, the term "oligonucleotide" refers to a nucleic acid comprising about 8 to about 50 covalently linked nucleotides, often comprising from about 8 to about 35 nucleotides, and more often from about 10 to about 25 nucleotides. The backbone and nucleotides within an oligonucleotide may be the same as those of naturally occurring nucleic acids, or analogs or derivatives of naturally occurring nucleic acids, provided that oligonucleotides having such analogs or derivatives retain the ability to hybridize specifically to a nucleic acid comprising a targeted polymorphism. Oligonucleotides described herein may be used as hybridization probes or as components of prognostic or diagnostic assays, for example, as described herein.

[0039] Oligonucleotides are typically synthesized using standard methods and equipment, such as the ABI™3900 High Throughput DNA Synthesizer and the EXPEDITE™ 8909 Nucleic Acid Synthesizer, both of which are available from Applied Biosystems (Foster City, CA). Analogs and derivatives are exemplified in U.S. Pat. Nos. 4,469,863; 5,536,821; 5,541,306; 5,637,683; 5,637,684;

5,700,922; 5,717,083; 5,719,262; 5,739,308; 5,773,601; 5,886,165; 5,929,226; 5,977,296; 6,140,482; WO 00/56746; WO 01/14398, and related publications. Methods for synthesizing oligonucleotides comprising such analogs or derivatives are disclosed, for example, in the patent publications cited above and in U.S. Pat. Nos. 5,614,622; 5,739,314; 5,955,599; 5,962,674; 6,117,992; in WO 00/75372; and in related publications.

[0040] Oligonucleotides may also be linked to a second moiety. The second moiety may be an additional nucleotide sequence such as a tail sequence (*e.g.*, a polyadenosine tail), an adapter sequence (*e.g.*, phage M13 universal tail sequence), and others. Alternatively, the second moiety may be a non-nucleotide moiety such as a moiety which facilitates linkage to a solid support or a label to facilitate detection of the oligonucleotide. Such labels include, without limitation, a radioactive label, a fluorescent label, a chemiluminescent label, a paramagnetic label, and the like. The second moiety may be attached to any position of the oligonucleotide, provided the oligonucleotide can hybridize to the nucleic acid comprising the polymorphism.

Uses for Nucleic Acid Sequence

[0041] Nucleic acid coding sequences may be used for diagnostic purposes for detection and control of polypeptide expression. Also, included herein are oligonucleotide sequences such as antisense RNA, small-interfering RNA (siRNA) and DNA molecules and ribozymes that function to inhibit translation of a polypeptide. Antisense techniques and RNA interference techniques are known in the art and are described herein.

[0042] Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, hammerhead motif ribozyme molecules may be engineered that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences corresponding to or complementary to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between fifteen (15) and twenty (20) ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for predicted structural features such as secondary structure that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

[0043] Antisense RNA and DNA molecules, siRNA and ribozymes may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically

synthesizing oligodeoxyribonucleotides well known in the art such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

[0044] DNA encoding a polypeptide also may have a number of uses for the diagnosis of diseases, including osteoarthritis, resulting from aberrant expression of a target gene described herein. For example, the nucleic acid sequence may be used in hybridization assays of biopsies or autopsies to diagnose abnormalities of expression or function (*e.g.*, Southern or Northern blot analysis, *in situ* hybridization assays).

[0045] In addition, the expression of a polypeptide during embryonic development may also be determined using nucleic acid encoding the polypeptide. As addressed, *infra*, production of functionally impaired polypeptide is the cause of various disease states, such as osteoarthritis. *In situ* hybridizations using polypeptide as a probe may be employed to predict problems related to osteoarthritis. Further, as indicated, *infra*, administration of human active polypeptide, recombinantly produced as described herein, may be used to treat disease states related to functionally impaired polypeptide. Alternatively, gene therapy approaches may be employed to remedy deficiencies of functional polypeptide or to replace or compete with dysfunctional polypeptide.

Expression Vectors, Host Cells, and Genetically Engineered Cells

[0046] Provided herein are nucleic acid vectors, often expression vectors, which contain a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence, or a substantially identical sequence thereof. As used herein, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked and can include a plasmid, cosmid, or viral vector. The vector can be capable of autonomous replication or it can integrate into a host DNA. Viral vectors may include replication defective retroviruses, adenoviruses and adeno-associated viruses for example.

[0047] A vector can include a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence in a form suitable for expression of an encoded target polypeptide or target nucleic acid in a host cell. A “target polypeptide” is a polypeptide encoded by a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence, or a substantially identical nucleotide sequence thereof. The recombinant expression vector typically includes one or more regulatory sequences operatively linked to the nucleic acid sequence to be expressed. The term “regulatory sequence” includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence, as well as tissue-specific regulatory and/or

inducible sequences. The design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, and the like. Expression vectors can be introduced into host cells to produce target polypeptides, including fusion polypeptides.

[0048] Recombinant expression vectors can be designed for expression of target polypeptides in prokaryotic or eukaryotic cells. For example, target polypeptides can be expressed in *E. coli*, insect cells (e.g., using baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

[0049] Expression of polypeptides in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polypeptides. Fusion vectors add a number of amino acids to a polypeptide encoded therein, usually to the amino terminus of the recombinant polypeptide. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant polypeptide; 2) to increase the solubility of the recombinant polypeptide; and 3) to aid in the purification of the recombinant polypeptide by acting as a ligand in affinity purification. Often, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant polypeptide to enable separation of the recombinant polypeptide from the fusion moiety subsequent to purification of the fusion polypeptide. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith & Johnson, *Gene* 67: 31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding polypeptide, or polypeptide A, respectively, to the target recombinant polypeptide.

[0050] Purified fusion polypeptides can be used in screening assays and to generate antibodies specific for target polypeptides. In a therapeutic embodiment, fusion polypeptide expressed in a retroviral expression vector is used to infect bone marrow cells that are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g., six (6) weeks).

[0051] Expressing the polypeptide in host bacteria with an impaired capacity to proteolytically cleave the recombinant polypeptide is often used to maximize recombinant polypeptide expression (Gottesman, S., *Gene Expression Technology: Methods in Enzymology*, Academic Press, San Diego, California 185: 119-128 (1990)). Another strategy is to alter the nucleotide sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, *Nucleic Acids Res.* 20: 2111-2118 (1992)). Such alteration of nucleotide sequences can be carried out by standard DNA synthesis techniques.

[0052] When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. Recombinant mammalian expression vectors are often capable of directing expression of the nucleic acid in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Non-limiting examples of suitable tissue-specific promoters include an albumin promoter (liver-specific; Pinkert *et al.*, *Genes Dev.* 1: 268-277 (1987)), lymphoid-specific promoters (Calame & Eaton, *Adv. Immunol.* 43: 235-275 (1988)), promoters of T cell receptors (Winoto & Baltimore, *EMBO J.* 8: 729-733 (1989)) promoters of immunoglobulins (Banerji *et al.*, *Cell* 33: 729-740 (1983); Queen & Baltimore, *Cell* 33: 741-748 (1983)), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne & Ruddle, *Proc. Natl. Acad. Sci. USA* 86: 5473-5477 (1989)), pancreas-specific promoters (Edlund *et al.*, *Science* 230: 912-916 (1985)), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are sometimes utilized, for example, the murine hox promoters (Kessel & Gruss, *Science* 249: 374-379 (1990)) and the α -fetoprotein promoter (Campes & Tilghman, *Genes Dev.* 3: 537-546 (1989)).

[0053] A *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid also may be cloned into an expression vector in an antisense orientation. Regulatory sequences (*e.g.*, viral promoters and/or enhancers) operatively linked to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid cloned in the antisense orientation can be chosen for directing constitutive, tissue specific or cell type specific expression of antisense RNA in a variety of cell types. Antisense expression vectors can be in the form of a recombinant plasmid, phagemid or attenuated virus. For a discussion of the regulation of gene expression using antisense genes *see, e.g.*, Weintraub *et al.*, *Antisense RNA as a molecular tool for genetic analysis, Reviews - Trends in Genetics*, Vol. 1(1) (1986).

[0054] Also provided herein are host cells that include a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence within a recombinant expression vector or a fragment of such a nucleotide sequence which facilitate homologous recombination into a specific site of the host cell genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. Such terms refer not only to the particular subject cell but rather also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell can be any prokaryotic or eukaryotic cell. For example, a target polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

[0055] Vectors can be introduced into host cells via conventional transformation or transfection techniques. As used herein, the terms “transformation” and “transfection” are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, transduction/infection, DEAE-dextran-mediated transfection, lipofection, or electroporation.

[0056] A host cell provided herein can be used to produce (*i.e.*, express) a target polypeptide or a substantially identical polypeptide thereof. Accordingly, further provided are methods for producing a target polypeptide using host cells described herein. In one embodiment, the method includes culturing host cells into which a recombinant expression vector encoding a target polypeptide has been introduced in a suitable medium such that a target polypeptide is produced. In another embodiment, the method further includes isolating a target polypeptide from the medium or the host cell.

[0057] Also provided are cells or purified preparations of cells which include a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* transgene, or which otherwise misexpress target polypeptide. Cell preparations can consist of human or non-human cells, *e.g.*, rodent cells, *e.g.*, mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* transgene (*e.g.*, a heterologous form of a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* gene, such as a human gene expressed in non-human cells). The transgene can be misexpressed, *e.g.*, overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene which misexpress an endogenous target polypeptide (*e.g.*, expression of a gene is disrupted, also known as a knockout). Such cells can serve as a model for studying disorders which are related to mutated or mis-expressed alleles or for use in drug screening. Also provided are human cells (*e.g.*, a hematopoietic stem cells) transfected with a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid.

[0058] Also provided are cells or a purified preparation thereof (*e.g.*, human cells) in which an endogenous *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid is under the control of a regulatory sequence that does not normally control the expression of the endogenous gene. The expression characteristics of an endogenous gene within a cell (*e.g.*, a cell line or microorganism) can be modified by inserting a heterologous DNA regulatory element into the genome of the cell such that the inserted regulatory element is operably linked to the corresponding endogenous gene. For example, an endogenous corresponding gene (*e.g.*, a gene which is “transcriptionally silent,” not normally expressed, or expressed only at very low levels) may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell. Techniques such as targeted homologous recombinations, can be used to insert the heterologous DNA as described in, *e.g.*, Chappel, US 5,272,071; WO 91/06667, published on May 16, 1991.

Transgenic Animals

[0059] Non-human transgenic animals that express a heterologous target polypeptide (*e.g.*, expressed from a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid or substantially identical sequence thereof) can be generated. Such animals are useful for studying the function and/or activity of a target polypeptide and for identifying and/or evaluating modulators of the activity of a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid or encoded polypeptide. As used herein, a “transgenic animal” is a non-human animal such as a mammal (*e.g.*, a non-human primate such as chimpanzee, baboon, or macaque; an ungulate such as an equine, bovine, or caprine; or a rodent such as a rat, a mouse, or an Israeli sand rat), a bird (*e.g.*, a chicken or a turkey), an amphibian (*e.g.*, a frog, salamander, or newt), or an insect (*e.g.*, *Drosophila melanogaster*), in which one or more of the cells of the animal includes a transgene. A transgene is exogenous DNA or a rearrangement (*e.g.*, a deletion of endogenous chromosomal DNA) that is often integrated into or occurs in the genome of cells in a transgenic animal. A transgene can direct expression of an encoded gene product in one or more cell types or tissues of the transgenic animal, and other transgenes can reduce expression (*e.g.*, a knockout). Thus, a transgenic animal can be one in which an endogenous nucleic acid homologous to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal (*e.g.*, an embryonic cell of the animal) prior to development of the animal.

[0060] Intronic sequences and polyadenylation signals can also be included in the transgene to increase expression efficiency of the transgene. One or more tissue-specific regulatory sequences can be operably linked to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence to direct expression of an encoded polypeptide to particular cells. A transgenic founder animal can be identified based upon the presence of a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence in its genome and/or expression of encoded mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence can further be bred to other transgenic animals carrying other transgenes.

[0061] Target polypeptides can be expressed in transgenic animals or plants by introducing, for example, a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid into the genome of an animal that encodes the target polypeptide. In preferred embodiments the nucleic acid is placed under the control of a tissue specific promoter, *e.g.*, a milk or egg specific promoter, and recovered from the milk or eggs produced by the animal. Also included is a population of cells from a transgenic animal.

Target Polypeptides

[0062] Also featured herein are isolated target polypeptides, which are encoded by a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence (e.g., SEQ ID NO: 1-12), or a substantially identical nucleotide sequence thereof. Examples of *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* polypeptides are set forth in SEQ ID NO: 13-20. The term “polypeptide” as used herein includes proteins and peptides. An “isolated” or “purified” polypeptide or protein is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. In one embodiment, the language “substantially free” means preparation of a target polypeptide having less than about 30%, 20%, 10% and more preferably 5% (by dry weight), of non-target polypeptide (also referred to herein as a “contaminating protein”), or of chemical precursors or non-target chemicals. When the target polypeptide or a biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, specifically, where culture medium represents less than about 20%, sometimes less than about 10%, and often less than about 5% of the volume of the polypeptide preparation. Isolated or purified target polypeptide preparations are sometimes 0.01 milligrams or more or 0.1 milligrams or more, and often 1.0 milligrams or more and 10 milligrams or more in dry weight.

[0063] Further included herein are target polypeptide fragments. The polypeptide fragment may be a domain or part of a domain of a target polypeptide. The polypeptide fragment may have increased, decreased or unexpected biological activity. The polypeptide fragment is often 50 or fewer, 100 or fewer, or 200 or fewer amino acids in length, and is sometimes 300, 400, 500, 600, 700, or 900 or fewer amino acids in length. Certain embodiments are directed to *PSMB1* polypeptide fragments (e.g., sequence accessed by NP_002784; rs756519 in Table A), such as a proteasome protease domain (e.g., starting at about amino acid 34 and ending at about amino acid 226) or a proteasome B domain (e.g., starting at about amino acid 41 and ending at about amino acid 88).

[0064] Substantially identical target polypeptides may depart from the amino acid sequences of target polypeptides in different manners. For example, conservative amino acid modifications may be introduced at one or more positions in the amino acid sequences of target polypeptides. A “conservative amino acid substitution” is one in which the amino acid is replaced by another amino acid having a similar structure and/or chemical function. Families of amino acid residues having similar structures and functions are well known. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine,

tryptophan, histidine). Also, essential and non-essential amino acids may be replaced. A “non-essential” amino acid is one that can be altered without abolishing or substantially altering the biological function of a target polypeptide, whereas altering an “essential” amino acid abolishes or substantially alters the biological function of a target polypeptide. Amino acids that are conserved among target polypeptides are typically essential amino acids. In certain embodiments, the polypeptide includes one or more non-synonymous polymorphic variants associated with osteoarthritis.

[0065] Also, target polypeptides may exist as chimeric or fusion polypeptides. As used herein, a target “chimeric polypeptide” or target “fusion polypeptide” includes a target polypeptide linked to a non-target polypeptide. A “non-target polypeptide” refers to a polypeptide having an amino acid sequence corresponding to a polypeptide which is not substantially identical to the target polypeptide, which includes, for example, a polypeptide that is different from the target polypeptide and derived from the same or a different organism. The target polypeptide in the fusion polypeptide can correspond to an entire or nearly entire target polypeptide or a fragment thereof. The non-target polypeptide can be fused to the N-terminus or C-terminus of the target polypeptide.

[0066] Fusion polypeptides can include a moiety having high affinity for a ligand. For example, the fusion polypeptide can be a GST-target fusion polypeptide in which the target sequences are fused to the C-terminus of the GST sequences, or a polyhistidine-target fusion polypeptide in which the target polypeptide is fused at the N- or C-terminus to a string of histidine residues. Such fusion polypeptides can facilitate purification of recombinant target polypeptide. Expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide), and a nucleotide sequence in SEQ ID NO: 1-12, or a substantially identical nucleotide sequence thereof, can be cloned into an expression vector such that the fusion moiety is linked in-frame to the target polypeptide. Further, the fusion polypeptide can be a target polypeptide containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression, secretion, cellular internalization, and cellular localization of a target polypeptide can be increased through use of a heterologous signal sequence. Fusion polypeptides can also include all or a part of a serum polypeptide (*e.g.*, an IgG constant region or human serum albumin).

[0067] Target polypeptides can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. Administration of these target polypeptides can be used to affect the bioavailability of a substrate of the target polypeptide and may effectively increase target polypeptide biological activity in a cell. Target fusion polypeptides may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a target polypeptide; (ii) mis-regulation of the gene encoding the target polypeptide; and (iii) aberrant post-translational modification of a target polypeptide. Also, target polypeptides can be used as immunogens to produce anti-target

antibodies in a subject, to purify target polypeptide ligands or binding partners, and in screening assays to identify molecules which inhibit or enhance the interaction of a target polypeptide with a substrate.

[0068] In addition, polypeptides can be chemically synthesized using techniques known in the art (See, *e.g.*, Creighton, 1983 *Proteins*. New York, N.Y.: W. H. Freeman and Company; and Hunkapiller et al., (1984) *Nature* July 12 -18;310(5973):105-11). For example, a relative short fragment can be synthesized by use of a peptide synthesizer. Furthermore, if desired, non-classical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the fragment sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid, α -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, γ -Abu, ϵ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoroamino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

[0069] Polypeptides and polypeptide fragments sometimes are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH₄; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; and the like. Additional post-translational modifications include, for example, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The polypeptide fragments may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the polypeptide.

[0070] Also provided are chemically modified derivatives of polypeptides that can provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (*see e.g.*, U.S. Pat. No: 4,179,337. The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

[0071] The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term

“about” indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

[0072] The polymers should be attached to the polypeptide with consideration of effects on functional or antigenic domains of the polypeptide. There are a number of attachment methods available to those skilled in the art (e.g., EP 0 401 384 (coupling PEG to G-CSF) and Malik et al. (1992) Exp Hematol. September;20(8):1028-35 (pegylation of GM-CSF using tresyl chloride)). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues, glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. For therapeutic purposes, the attachment sometimes is at an amino group, such as attachment at the N-terminus or lysine group.

[0073] Proteins can be chemically modified at the N-terminus. Using polyethylene glycol as an illustration of such a composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, and the like), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus may be accomplished by reductive alkylation, which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

Substantially Identical Nucleic Acids and Polypeptides

[0074] Nucleotide sequences and polypeptide sequences that are substantially identical to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence and the target polypeptide sequences encoded by those nucleotide sequences, respectively, are included herein. The term “substantially identical” as used herein refers to two or more nucleic acids or polypeptides sharing one or more identical nucleotide sequences or polypeptide sequences, respectively. Included are nucleotide sequences or polypeptide

sequences that are 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more (each often within a 1%, 2%, 3% or 4% variability) identical to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence or the encoded target polypeptide amino acid sequences. One test for determining whether two nucleic acids are substantially identical is to determine the percent of identical nucleotide sequences or polypeptide sequences shared between the nucleic acids or polypeptides.

[0075] Calculations of sequence identity are often performed as follows. Sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The length of a reference sequence aligned for comparison purposes is sometimes 30% or more, 40% or more, 50% or more, often 60% or more, and more often 70% or more, 80% or more, 90% or more, or 100% of the length of the reference sequence. The nucleotides or amino acids at corresponding nucleotide or polypeptide positions, respectively, are then compared among the two sequences. When a position in the first sequence is occupied by the same nucleotide or amino acid as the corresponding position in the second sequence, the nucleotides or amino acids are deemed to be identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, introduced for optimal alignment of the two sequences.

[0076] Comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. Percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of Meyers & Miller, *CABIOS* 4: 11-17 (1989), which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. Also, percent identity between two amino acid sequences can be determined using the Needleman & Wunsch, *J. Mol. Biol.* 48: 444-453 (1970) algorithm which has been incorporated into the GAP program in the GCG software package (available at the http address www.gcg.com), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. Percent identity between two nucleotide sequences can be determined using the GAP program in the GCG software package (available at http address www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A set of parameters often used is a Blossum 62 scoring matrix with a gap open penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

[0077] Another manner for determining if two nucleic acids are substantially identical is to assess whether a polynucleotide homologous to one nucleic acid will hybridize to the other nucleic acid under stringent conditions. As used herein, the term "stringent conditions" refers to conditions for hybridization and washing. Stringent conditions are known to those skilled in the art and can be found in *Current*

Protocols in Molecular Biology, John Wiley & Sons, N.Y. , 6.3.1-6.3.6 (1989). Aqueous and non-aqueous methods are described in that reference and either can be used. An example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50°C. Another example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 55°C. A further example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C. Often, stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 65°C. More often, stringency conditions are 0.5M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C.

[0078] An example of a substantially identical nucleotide sequence to a nucleotide sequence in SEQ ID NO: 1-12 is one that has a different nucleotide sequence but still encodes the same polypeptide sequence encoded by the nucleotide sequence in SEQ ID NO: 1-12. Another example is a nucleotide sequence that encodes a polypeptide having a polypeptide sequence that is more than 70% or more identical to, sometimes more than 75% or more, 80% or more, or 85% or more identical to, and often more than 90% or more and 95% or more identical to a polypeptide sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12.

[0079] Nucleotide sequences in SEQ ID NO: 1-12 and amino acid sequences of encoded polypeptides can be used as “query sequences” to perform a search against public databases to identify other family members or related sequences, for example. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul *et al.*, *J. Mol. Biol.* 215: 403-10 (1990). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleotide sequences in SEQ ID NO: 1-12. BLAST polypeptide searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to polypeptides encoded by the nucleotide sequences of SEQ ID NO: 1-12. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, *Nucleic Acids Res.* 25(17): 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used (*see* the http address www.ncbi.nlm.nih.gov).

[0080] A nucleic acid that is substantially identical to a nucleotide sequence in SEQ ID NO: 1 may include polymorphic sites at positions equivalent to those described herein when the sequences are aligned. For example, using the alignment procedures described herein, SNPs in a sequence substantially identical to a sequence in SEQ ID NO: 1-12 can be identified at nucleotide positions that match (*i.e.*,

align) with nucleotides at SNP positions in each nucleotide sequence in SEQ ID NO: 1-12. Also, where a polymorphic variation results in an insertion or deletion, insertion or deletion of a nucleotide sequence from a reference sequence can change the relative positions of other polymorphic sites in the nucleotide sequence.

[0081] Substantially identical nucleotide and polypeptide sequences include those that are naturally occurring, such as allelic variants (same locus), splice variants, homologs (different locus), and orthologs (different organism) or can be non-naturally occurring. Non-naturally occurring variants can be generated by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product). Orthologs, homologs, allelic variants, and splice variants can be identified using methods known in the art. These variants normally comprise a nucleotide sequence encoding a polypeptide that is 50% or more, about 55% or more, often about 70-75% or more or about 80-85% or more, and sometimes about 90-95% or more identical to the amino acid sequences of target polypeptides or a fragment thereof. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions to a nucleotide sequence in SEQ ID NO: 1-12 or a fragment of this sequence. Nucleic acid molecules corresponding to orthologs, homologs, and allelic variants of a nucleotide sequence in SEQ ID NO: 1-12 can further be identified by mapping the sequence to the same chromosome or locus as the nucleotide sequence in SEQ ID NO: 1-12.

[0082] Also, substantially identical nucleotide sequences may include codons that are altered with respect to the naturally occurring sequence for enhancing expression of a target polypeptide in a particular expression system. For example, the nucleic acid can be one in which one or more codons are altered, and often 10% or more or 20% or more of the codons are altered for optimized expression in bacteria (*e.g.*, *E. coli*), yeast (*e.g.*, *S. cerevisiae*), human (*e.g.*, 293 cells), insect, or rodent (*e.g.*, hamster) cells.

Methods for Identifying Risk of osteoarthritis

[0083] Methods for prognosing and diagnosing osteoarthritis are included herein. These methods include detecting the presence or absence of one or more polymorphic variations in a nucleotide sequence associated with osteoarthritis, such as variants in or around the loci set forth herein, or a substantially identical sequence thereof, in a sample from a subject, where the presence of a polymorphic variant described herein is indicative of a risk of osteoarthritis. Determining a risk of osteoarthritis sometimes refers to determining whether an individual is at an increased risk of osteoarthritis (*e.g.*, intermediate risk or higher risk).

[0084] Thus, featured herein is a method for identifying a subject who is at risk of osteoarthritis, which comprises detecting an aberration associated with osteoarthritis in a nucleic acid sample from the subject. An embodiment is a method for detecting a risk of osteoarthritis in a subject, which comprises detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-12; (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-12; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-12, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-12; and (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic site; whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject. In certain embodiments, polymorphic variants at the positions described herein are detected for determining a risk of osteoarthritis, and polymorphic variants at positions in linkage disequilibrium with these positions are detected for determining a risk of osteoarthritis. As used herein, "SEQ ID NO: 1-12" refers to individual sequences of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12, each sequence being separately applicable to embodiments described herein.

[0085] Risk of osteoarthritis sometimes is expressed as a probability, such as an odds ratio, percentage, or risk factor. Risk often is based upon the presence or absence of one or more polymorphic variants described herein, and also may be based in part upon phenotypic traits of the individual being tested. Methods for calculating risk based upon patient data are well known (*see, e.g.,* Agresti, *Categorical Data Analysis*, 2nd Ed. 2002. Wiley). Allelotyping and genotyping analyses may be carried out in populations other than those exemplified herein to enhance the predictive power of the prognostic method. These further analyses are executed in view of the exemplified procedures described herein, and may be based upon the same polymorphic variations or additional polymorphic variations.

[0086] In certain embodiments, determining the presence of a combination of two or more polymorphic variants associated with osteoarthritis in one or more genetic loci (e.g., one or more genes) of the sample is determined to identify, quantify and/or estimate, risk of osteoarthritis. The risk often is the probability of having or developing osteoarthritis. The risk sometimes is expressed as a relative risk with respect to a population average risk of osteoarthritis, and sometimes is expressed as a relative risk with respect to the lowest risk group. Such relative risk assessments often are based upon penetrance values determined by statistical methods, and are particularly useful to clinicians and insurance companies for assessing risk of osteoarthritis (e.g., a clinician can target appropriate detection, prevention and therapeutic regimens to a patient after determining the patient's risk of osteoarthritis, and an insurance company can fine tune actuarial tables based upon population genotype assessments of

osteoarthritis risk). Risk of osteoarthritis sometimes is expressed as an odds ratio, which is the odds of a particular person having a genotype has or will develop osteoarthritis with respect to another genotype group (e.g., the most disease protective genotype or population average). In related embodiments, the determination is utilized to identify a subject at risk of osteoarthritis. In an embodiment, two or more polymorphic variations are detected in two or more regions in human genomic DNA associated with increased risk of osteoarthritis, such as a locus containing a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG*, for example. In certain embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected in the sample. In specific embodiments, polymorphic variants are detected in a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* region, for example. In another embodiment, polymorphic variants are detected at two or more positions selected from the group consisting of rs756519, rs1042327, rs8770, rs1563055, rs912428 and rs1888475; 229, 16559, 44850 and/or 46252 in SEQ ID NO: 1; 40985, 46168, 51925 and/or 52168 in SEQ ID NO: 2; 49693, 82475, 92462 and/or 96275 in SEQ ID NO: 3; and 33137, 55499 and/or 58282 in SEQ ID NO: 4. In certain embodiments, polymorphic variants are detected at other genetic loci (e.g., the polymorphic variants can be detected in *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* in addition to other loci or only in other loci), where the other loci include but are not limited to those described in concurrently-filed patent applications having attorney docket number 524593008800, 524593008900, 524593009000, 524593009100 or 524593009200, each of which is incorporated herein by reference in its entirety.

[0087] Results from prognostic tests may be combined with other test results to diagnose osteoarthritis. For example, prognostic results may be gathered, a patient sample may be ordered based on a determined predisposition to osteoarthritis, the patient sample is analyzed, and the results of the analysis may be utilized to diagnose osteoarthritis. Also osteoarthritis diagnostic method can be developed from studies used to generate prognostic methods in which populations are stratified into subpopulations having different progressions of osteoarthritis. In another embodiment, prognostic results may be gathered, a patient's risk factors for developing osteoarthritis (e.g., age, weight, race, diet) analyzed, and a patient sample may be ordered based on a determined predisposition to osteoarthritis.

[0088] The nucleic acid sample typically is isolated from a biological sample obtained from a subject. For example, nucleic acid can be isolated from blood, saliva, sputum, urine, cell scrapings, and biopsy tissue. The nucleic acid sample can be isolated from a biological sample using standard techniques, such as the technique described in Example 2. As used herein, the term "subject" refers primarily to humans but also refers to other mammals such as dogs, cats, and ungulates (e.g., cattle, sheep, and swine). Subjects also include avians (e.g., chickens and turkeys), reptiles, and fish (e.g., salmon), as embodiments described herein can be adapted to nucleic acid samples isolated from any of these organisms. The nucleic acid sample may be isolated from the subject and then directly utilized in a

method for determining the presence of a polymorphic variant, or alternatively, the sample may be isolated and then stored (*e.g.*, frozen) for a period of time before being subjected to analysis.

[0089] The presence or absence of a polymorphic variant is determined using one or both chromosomal complements represented in the nucleic acid sample. Determining the presence or absence of a polymorphic variant in both chromosomal complements represented in a nucleic acid sample from a subject having a copy of each chromosome is useful for determining the zygosity of an individual for the polymorphic variant (*i.e.*, whether the individual is homozygous or heterozygous for the polymorphic variant). Any oligonucleotide-based diagnostic may be utilized to determine whether a sample includes the presence or absence of a polymorphic variant in a sample. For example, primer extension methods, ligase sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,679,524 and 5,952,174, and WO 01/27326), mismatch sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,851,770; 5,958,692; 6,110,684; and 6,183,958), microarray sequence determination methods, restriction fragment length polymorphism (RFLP), single strand conformation polymorphism detection (SSCP) (*e.g.*, U.S. Pat. Nos. 5,891,625 and 6,013,499), PCR-based assays (*e.g.*, TAQMAN[®] PCR System (Applied Biosystems)), and nucleotide sequencing methods may be used.

[0090] Oligonucleotide extension methods typically involve providing a pair of oligonucleotide primers in a polymerase chain reaction (PCR) or in other nucleic acid amplification methods for the purpose of amplifying a region from the nucleic acid sample that comprises the polymorphic variation. One oligonucleotide primer is complementary to a region 3' of the polymorphism and the other is complementary to a region 5' of the polymorphism. A PCR primer pair may be used in methods disclosed in U.S. Pat. Nos. 4,683,195; 4,683,202, 4,965,188; 5,656,493; 5,998,143; 6,140,054; WO 01/27327; and WO 01/27329 for example. PCR primer pairs may also be used in any commercially available machines that perform PCR, such as any of the GENEAMP[®] Systems available from Applied Biosystems. Also, those of ordinary skill in the art will be able to design oligonucleotide primers based upon a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence using knowledge available in the art.

[0091] Also provided is an extension oligonucleotide that hybridizes to the amplified fragment adjacent to the polymorphic variation. As used herein, the term "adjacent" refers to the 3' end of the extension oligonucleotide being often 1 nucleotide from the 5' end of the polymorphic site, and sometimes 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5' end of the polymorphic site, in the nucleic acid when the extension oligonucleotide is hybridized to the nucleic acid. The extension oligonucleotide then is extended by one or more nucleotides, and the number and/or type of nucleotides that are added to the extension oligonucleotide determine whether the polymorphic variant is present. Oligonucleotide extension methods are disclosed, for example, in U.S. Pat. Nos. 4,656,127; 4,851,331; 5,679,524; 5,834,189; 5,876,934; 5,908,755; 5,912,118; 5,976,802; 5,981,186; 6,004,744; 6,013,431; 6,017,702;

6,046,005; 6,087,095; 6,210,891; and WO 01/20039. Oligonucleotide extension methods using mass spectrometry are described, for example, in U.S. Pat. Nos. 5,547,835; 5,605,798; 5,691,141; 5,849,542; 5,869,242; 5,928,906; 6,043,031; and 6,194,144, and a method often utilized is described herein in Example 2.

[0092] A microarray can be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A microarray may include any oligonucleotides described herein, and methods for making and using oligonucleotide microarrays suitable for diagnostic use are disclosed in U.S. Pat. Nos. 5,492,806; 5,525,464; 5,589,330; 5,695,940; 5,849,483; 6,018,041; 6,045,996; 6,136,541; 6,142,681; 6,156,501; 6,197,506; 6,223,127; 6,225,625; 6,229,911; 6,239,273; WO 00/52625; WO 01/25485; and WO 01/29259. The microarray typically comprises a solid support and the oligonucleotides may be linked to this solid support by covalent bonds or by non-covalent interactions. The oligonucleotides may also be linked to the solid support directly or by a spacer molecule. A microarray may comprise one or more oligonucleotides complementary to a polymorphic site set forth herein.

[0093] A kit also may be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A kit often comprises one or more pairs of oligonucleotide primers useful for amplifying a fragment of a nucleotide sequence of SEQ ID NO: 1-12 or a substantially identical sequence thereof, where the fragment includes a polymorphic site. The kit sometimes comprises a polymerizing agent, for example, a thermostable nucleic acid polymerase such as one disclosed in U.S. Pat. Nos. 4,889,818 or 6,077,664. Also, the kit often comprises an elongation oligonucleotide that hybridizes to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence in a nucleic acid sample adjacent to the polymorphic site. Where the kit includes an elongation oligonucleotide, it also often comprises chain elongating nucleotides, such as dATP, dTTP, dGTP, dCTP, and dITP, including analogs of dATP, dTTP, dGTP, dCTP and dITP, provided that such analogs are substrates for a thermostable nucleic acid polymerase and can be incorporated into a nucleic acid chain elongated from the extension oligonucleotide. Along with chain elongating nucleotides would be one or more chain terminating nucleotides such as ddATP, ddTTP, ddGTP, ddCTP, and the like. In an embodiment, the kit comprises one or more oligonucleotide primer pairs, a polymerizing agent, chain elongating nucleotides, at least one elongation oligonucleotide, and one or more chain terminating nucleotides. Kits optionally include buffers, vials, microtiter plates, and instructions for use.

[0094] An individual identified as being at risk of osteoarthritis may be heterozygous or homozygous with respect to the allele associated with a higher risk of osteoarthritis. A subject homozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively high risk of osteoarthritis, a subject heterozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively intermediate risk of osteoarthritis, and a subject homozygous for an

allele associated with a decreased risk of osteoarthritis is at a comparatively low risk of osteoarthritis. A genotype may be assessed for a complementary strand, such that the complementary nucleotide at a particular position is detected.

[0095] Also featured are methods for determining risk of osteoarthritis and/or identifying a subject at risk of osteoarthritis by contacting a polypeptide or protein encoded by a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence from a subject with an antibody that specifically binds to an epitope associated with increased risk of osteoarthritis in the polypeptide.

Applications of Prognostic and Diagnostic Results to Pharmacogenomic Methods

[0096] Pharmacogenomics is a discipline that involves tailoring a treatment for a subject according to the subject's genotype as a particular treatment regimen may exert a differential effect depending upon the subject's genotype. For example, based upon the outcome of a prognostic test described herein, a clinician or physician may target pertinent information and preventative or therapeutic treatments to a subject who would be benefited by the information or treatment and avoid directing such information and treatments to a subject who would not be benefited (*e.g.*, the treatment has no therapeutic effect and/or the subject experiences adverse side effects).

[0097] The following is an example of a pharmacogenomic embodiment. A particular treatment regimen can exert a differential effect depending upon the subject's genotype. Where a candidate therapeutic exhibits a significant interaction with a major allele and a comparatively weak interaction with a minor allele (*e.g.*, an order of magnitude or greater difference in the interaction), such a therapeutic typically would not be administered to a subject genotyped as being homozygous for the minor allele, and sometimes not administered to a subject genotyped as being heterozygous for the minor allele. In another example, where a candidate therapeutic is not significantly toxic when administered to subjects who are homozygous for a major allele but is comparatively toxic when administered to subjects heterozygous or homozygous for a minor allele, the candidate therapeutic is not typically administered to subjects who are genotyped as being heterozygous or homozygous with respect to the minor allele.

[0098] The methods described herein are applicable to pharmacogenomic methods for preventing, alleviating or treating osteoarthritis. For example, a nucleic acid sample from an individual may be subjected to a prognostic test described herein. Where one or more polymorphic variations associated with increased risk of osteoarthritis are identified in a subject, information for preventing or treating osteoarthritis and/or one or more osteoarthritis treatment regimens then may be prescribed to that subject.

[0099] In certain embodiments, a treatment or preventative regimen is specifically prescribed and/or administered to individuals who will most benefit from it based upon their risk of developing osteoarthritis assessed by the methods described herein. Thus, provided are methods for identifying a subject predisposed to osteoarthritis and then prescribing a therapeutic or preventative regimen to

individuals identified as having a predisposition. Thus, certain embodiments are directed to a method for reducing osteoarthritis in a subject, which comprises: detecting the presence or absence of a polymorphic variant associated with osteoarthritis in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-12; (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-12; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-12, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-12; and (d) a fragment of a polynucleotide sequence of (a), (b), or (c); and prescribing or administering a treatment regimen to a subject from whom the sample originated where the presence of a polymorphic variation associated with osteoarthritis is detected in the nucleotide sequence. In these methods, predisposition results may be utilized in combination with other test results to diagnose osteoarthritis.

[0100] Certain preventative treatments often are prescribed to subjects having a predisposition to osteoarthritis and where the subject is diagnosed with osteoarthritis or is diagnosed as having symptoms indicative of an early stage of osteoarthritis. The treatment sometimes is preventative (e.g., is prescribed or administered to reduce the probability that osteoarthritis arises or progresses), sometimes is therapeutic, and sometimes delays, alleviates or halts the progression of osteoarthritis. Any known preventative or therapeutic treatment for alleviating or preventing the occurrence of osteoarthritis is prescribed and/or administered. For example, the treatment often is directed to decreasing pain and improving joint movement. Examples of OA treatments include exercises to keep joints flexible and improve muscle strength. Different medications to control pain, including corticosteroids and nonsteroidal anti-inflammatory drugs (NSAIDs, e.g., Voltaren); cyclooxygenase-2 (COX-2) inhibitors (e.g., Celebrex, Vioxx, Mobic, and Bextra); monoclonal antibodies (e.g., Remicade); tumor necrosis factor inhibitors (e.g., Enbrel); or injections of glucocorticoids, hyaluronic acid or chondroitin sulfate into joints that are inflamed and not responsive to NSAIDs. Orally administered chondroitin sulfate also may be used as a therapeutic, as it may increase hyaluronic acid levels and viscosity of synovial fluid, and decrease collagenase levels in synovial fluid. Also, glucosamine can serve as an OA therapeutic as delivering it into joints may inhibit enzymes involved in cartilage degradation and enhance the production of hyaluronic acid. For mild pain without inflammation, acetaminophen may be used. Other treatments include: heat/cold therapy for temporary pain relief; joint protection to prevent strain or stress on painful joints; surgery to relieve chronic pain in damaged joints; and weight control to prevent extra stress on weight-bearing joints.

[0101] As therapeutic approaches for treating osteoarthritis continue to evolve and improve, the goal of treatments for osteoarthritis related disorders is to intervene even before clinical signs first manifest.

Thus, genetic markers associated with susceptibility to osteoarthritis prove useful for early diagnosis, prevention and treatment of osteoarthritis.

[0102] As osteoarthritis preventative and treatment information can be specifically targeted to subjects in need thereof (*e.g.*, those at risk of developing osteoarthritis or those in an early stage of osteoarthritis), provided herein is a method for preventing or reducing the risk of developing osteoarthritis in a subject, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying a subject with a predisposition to osteoarthritis, whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject; and (c) if such a predisposition is identified, providing the subject with information about methods or products to prevent or reduce osteoarthritis or to delay the onset of osteoarthritis. Also provided is a method of targeting information or advertising to a subpopulation of a human population based on the subpopulation being genetically predisposed to a disease or condition, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; and (c) providing information only to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition.

[0103] Pharmacogenomics methods also may be used to analyze and predict a response to osteoarthritis treatment or a drug. For example, if pharmacogenomics analysis indicates a likelihood that an individual will respond positively to osteoarthritis treatment with a particular drug, the drug may be administered to the individual. Conversely, if the analysis indicates that an individual is likely to respond negatively to treatment with a particular drug, an alternative course of treatment may be prescribed. A negative response may be defined as either the absence of an efficacious response or the presence of toxic side effects. The response to a therapeutic treatment can be predicted in a background study in which subjects in any of the following populations are genotyped: a population that responds favorably to a treatment regimen, a population that does not respond significantly to a treatment regimen, and a population that responds adversely to a treatment regimen (*e.g.*, exhibits one or more side effects). These populations are provided as examples and other populations and subpopulations may be analyzed. Based upon the results of these analyses, a subject is genotyped to predict whether he or she will respond favorably to a treatment regimen, not respond significantly to a treatment regimen, or respond adversely to a treatment regimen.

[0104] The tests described herein also are applicable to clinical drug trials. One or more polymorphic variants indicative of response to an agent for treating osteoarthritis or to side effects to an agent for treating osteoarthritis may be identified using the methods described herein. Thereafter,

potential participants in clinical trials of such an agent may be screened to identify those individuals most likely to respond favorably to the drug and exclude those likely to experience side effects. In that way, the effectiveness of drug treatment may be measured in individuals who respond positively to the drug, without lowering the measurement as a result of the inclusion of individuals who are unlikely to respond positively in the study and without risking undesirable safety problems.

[0105] Thus, another embodiment is a method of selecting an individual for inclusion in a clinical trial of a treatment or drug comprising the steps of: (a) obtaining a nucleic acid sample from an individual; (b) determining the identity of a polymorphic variation which is associated with a positive response to the treatment or the drug, or at least one polymorphic variation which is associated with a negative response to the treatment or the drug in the nucleic acid sample, and (c) including the individual in the clinical trial if the nucleic acid sample contains said polymorphic variation associated with a positive response to the treatment or the drug or if the nucleic acid sample lacks said polymorphic variation associated with a negative response to the treatment or the drug. In addition, the methods described herein for selecting an individual for inclusion in a clinical trial of a treatment or drug encompass methods with any further limitation described in this disclosure, or those following, specified alone or in any combination. The polymorphic variation may be in a sequence selected individually or in any combination from the group consisting of (i) a nucleotide sequence of SEQ ID NO: 1-12; (ii) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-12; (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-12, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-12; and (iv) a fragment of a polynucleotide sequence of (i), (ii), or (iii) comprising the polymorphic site. The including step (c) optionally comprises administering the drug or the treatment to the individual if the nucleic acid sample contains the polymorphic variation associated with a positive response to the treatment or the drug and the nucleic acid sample lacks said biallelic marker associated with a negative response to the treatment or the drug.

[0106] Also provided herein is a method of partnering between a diagnostic/prognostic testing provider and a provider of a consumable product, which comprises: (a) the diagnostic/prognostic testing provider detects the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) the diagnostic/prognostic testing provider identifies the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; (c) the diagnostic/prognostic testing provider forwards information to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition; and (d) the

provider of a consumable product forwards to the diagnostic test provider a fee every time the diagnostic/prognostic test provider forwards information to the subject as set forth in step (c) above.

Compositions Comprising Osteoarthritis-Directed Molecules

[0107] Featured herein is a composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and one or more molecules specifically directed and targeted to a nucleic acid comprising a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence or amino acid sequence. Such directed molecules include, but are not limited to, a compound that binds to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence or amino acid sequence referenced herein; a RNAi or siRNA molecule having a strand complementary or substantially complementary to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence (e.g., hybridizes to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence under conditions of high stringency); an antisense nucleic acid complementary or substantially complementary to an RNA encoded by a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence (e.g., hybridizes to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence under conditions of high stringency); a ribozyme that hybridizes to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence (e.g., hybridizes to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence under conditions of high stringency); a nucleic acid aptamer that specifically binds a polypeptide encoded by *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence; and an antibody that specifically binds to a polypeptide encoded by *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence or binds to a nucleic acid having such a nucleotide sequence. In specific embodiments, the osteoarthritis directed molecule interacts with a nucleic acid or polypeptide variant associated with osteoarthritis, such as variants referenced herein. In other embodiments, the osteoarthritis directed molecule interacts with a polypeptide involved in a signal pathway of a polypeptide encoded by a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence, or a nucleic acid comprising such a nucleotide sequence.

[0108] Compositions sometimes include an adjuvant known to stimulate an immune response, and in certain embodiments, an adjuvant that stimulates a T-cell lymphocyte response. Adjuvants are known, including but not limited to an aluminum adjuvant (e.g., aluminum hydroxide); a cytokine adjuvant or adjuvant that stimulates a cytokine response (e.g., interleukin (IL)-12 and/or gamma-interferon cytokines); a Freund-type mineral oil adjuvant emulsion (e.g., Freund's complete or incomplete adjuvant); a synthetic lipid compound; a copolymer adjuvant (e.g., TitreMax); a saponin; Quil A; a liposome; an oil-in-water emulsion (e.g., an emulsion stabilized by Tween 80 and pluronic polyoxyethylene/polyoxypropylene block copolymer (Syntex Adjuvant Formulation); TitreMax; detoxified endotoxin (MPL) and mycobacterial cell wall components (TDW, CWS) in 2% squalene (Ribi Adjuvant System)); a muramyl dipeptide; an immune-stimulating complex (ISCOM, e.g., an Ag-

modified saponin/cholesterol micelle that forms stable cage-like structure); an aqueous phase adjuvant that does not have a depot effect (e.g., Gerbu adjuvant); a carbohydrate polymer (e.g., AdjuPrime); L-tyrosine; a manide-oleate compound (e.g., Montanide); an ethylene-vinyl acetate copolymer (e.g., Elvax 40W1,2); or lipid A, for example. Such compositions are useful for generating an immune response against osteoarthritis directed molecule (e.g., an HLA-binding subsequence within a polypeptide encoded by a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence). In such methods, a peptide having an amino acid subsequence of a polypeptide encoded by a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence is delivered to a subject, where the subsequence binds to an HLA molecule and induces a CTL lymphocyte response. The peptide sometimes is delivered to the subject as an isolated peptide or as a minigene in a plasmid that encodes the peptide. Methods for identifying HLA-binding subsequences in such polypeptides are known (see e.g., publication WO02/20616 and PCT application US98/01373 for methods of identifying such sequences).

[0109] The cell may be in a group of cells cultured *in vitro* or in a tissue maintained *in vitro* or present in an animal *in vivo* (e.g., a rat, mouse, ape or human). In certain embodiments, a composition comprises a component from a cell such as a nucleic acid molecule (e.g., genomic DNA), a protein mixture or isolated protein, for example. The aforementioned compositions have utility in diagnostic, prognostic and pharmacogenomic methods described previously and in therapeutics described hereafter. Certain osteoarthritis directed molecules are described in greater detail below.

Compounds

[0110] Compounds can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive (see, e.g., Zuckermann et al., J. Med. Chem. 37: 2678-85 (1994)); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; "one-bead one-compound" library methods; and synthetic library methods using affinity chromatography selection. Biological library and peptoid library approaches are typically limited to peptide libraries, while the other approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, Anticancer Drug Des. 12: 145, (1997)). Examples of methods for synthesizing molecular libraries are described, for example, in DeWitt et al., Proc. Natl. Acad. Sci. U.S.A. 90: 6909 (1993); Erb et al., Proc. Natl. Acad. Sci. USA 91: 11422 (1994); Zuckermann et al., J. Med. Chem. 37: 2678 (1994); Cho et al., Science 261: 1303 (1993); Carrell et al., Angew. Chem. Int. Ed. Engl. 33: 2059 (1994); Carell et al., Angew. Chem. Int. Ed. Engl. 33: 2061 (1994); and in Gallop et al., J. Med. Chem. 37: 1233 (1994).

[0111] Libraries of compounds may be presented in solution (e.g., Houghten, *Biotechniques* 13: 412-421 (1992)), or on beads (Lam, *Nature* 354: 82-84 (1991)), chips (Fodor, *Nature* 364: 555-556 (1993)), bacteria or spores (Ladner, United States Patent No. 5,223,409), plasmids (Cull et al., *Proc. Natl. Acad. Sci. USA* 89: 1865-1869 (1992)) or on phage (Scott and Smith, *Science* 249: 386-390 (1990); Devlin, *Science* 249: 404-406 (1990); Cwirla et al., *Proc. Natl. Acad. Sci.* 87: 6378-6382 (1990); Felici, *J. Mol. Biol.* 222: 301-310 (1991); Ladner supra.).

[0112] A compound sometimes alters expression and sometimes alters activity of a polypeptide target and may be a small molecule. Small molecules include, but are not limited to, peptides, peptidomimetics (e.g., peptoids), amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

Antisense Nucleic Acid Molecules, Ribozymes, RNAi, siRNA and Modified Nucleic Acid Molecules

[0113] An “antisense” nucleic acid refers to a nucleotide sequence complementary to a “sense” nucleic acid encoding a polypeptide, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. The antisense nucleic acid can be complementary to an entire coding strand, or to a portion thereof or a substantially identical sequence thereof. In another embodiment, the antisense nucleic acid molecule is antisense to a “noncoding region” of the coding strand of a nucleotide sequence (e.g., 5’ and 3’ untranslated regions in SEQ ID NO: 1).

[0114] An antisense nucleic acid can be designed such that it is complementary to the entire coding region of an mRNA encoded by a nucleotide sequence (e.g., SEQ ID NO: 1), and often the antisense nucleic acid is an oligonucleotide antisense to only a portion of a coding or noncoding region of the mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of the mRNA, e.g., between the -10 and +10 regions of the target gene nucleotide sequence of interest. An antisense oligonucleotide can be, for example, about 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more nucleotides in length. The antisense nucleic acids, which include the ribozymes described hereafter, can be designed to target a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence, often a variant associated with osteoarthritis, or a substantially identical sequence thereof. Among the variants, minor alleles and major alleles can be targeted, and

those associated with a higher risk of osteoarthritis are often designed, tested, and administered to subjects.

[0115] An antisense nucleic acid can be constructed using chemical synthesis and enzymatic ligation reactions using standard procedures. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Antisense nucleic acid also can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

[0116] When utilized as therapeutics, antisense nucleic acids typically are administered to a subject (e.g., by direct injection at a tissue site) or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a polypeptide and thereby inhibit expression of the polypeptide, for example, by inhibiting transcription and/or translation. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then are administered systemically. For systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, for example, by linking antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. Antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. Sufficient intracellular concentrations of antisense molecules are achieved by incorporating a strong promoter, such as a pol II or pol III promoter, in the vector construct.

[0117] Antisense nucleic acid molecules sometimes are alpha-anomeric nucleic acid molecules. An alpha-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual beta-units, the strands run parallel to each other (Gaultier et al., Nucleic Acids. Res. 15: 6625-6641 (1987)). Antisense nucleic acid molecules can also comprise a 2'-O-methylribonucleotide (Inoue et al., Nucleic Acids Res. 15: 6131-6148 (1987)) or a chimeric RNA-DNA analogue (Inoue et al., FEBS Lett. 215: 327-330 (1987)). Antisense nucleic acids sometimes are composed of DNA or PNA or any other nucleic acid derivatives described previously.

[0118] In another embodiment, an antisense nucleic acid is a ribozyme. A ribozyme having specificity for a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence can include one or more sequences complementary to such a nucleotide sequence, and a sequence having a known catalytic region responsible for mRNA cleavage (see e.g., U.S. Pat. No. 5,093,246 or Haselhoff and Gerlach, Nature 334: 585-591 (1988)). For example, a derivative of a Tetrahymena L-19 IVS RNA is sometimes utilized in which the nucleotide sequence of the active site is complementary to the nucleotide sequence

to be cleaved in a mRNA (see e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742). Also, target mRNA sequences can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see e.g., Bartel & Szostak, Science 261: 1411-1418 (1993)).

[0119] Osteoarthritis directed molecules include in certain embodiments nucleic acids that can form triple helix structures with a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence, or a substantially identical sequence thereof, especially one that includes a regulatory region that controls expression of a polypeptide. Gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of a nucleotide sequence referenced herein or a substantially identical sequence (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of a gene in target cells (see e.g., Helene, Anticancer Drug Des. 6(6): 569-84 (1991); Helene et al., Ann. N.Y. Acad. Sci. 660: 27-36 (1992); and Maher, Bioassays 14(12): 807-15 (1992). Potential sequences that can be targeted for triple helix formation can be increased by creating a so-called “switchback” nucleic acid molecule. Switchback molecules are synthesized in an alternating 5’-3’, 3’-5’ manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

[0120] Osteoarthritis directed molecules include RNAi and siRNA nucleic acids. Gene expression may be inhibited by the introduction of double-stranded RNA (dsRNA), which induces potent and specific gene silencing, a phenomenon called RNA interference or RNAi. See, e.g., Fire et al., US Patent Number 6,506,559; Tuschl et al. PCT International Publication No. WO 01/75164; Kay et al. PCT International Publication No. WO 03/010180A1; or Bosher JM, Labouesse, Nat Cell Biol 2000 Feb;2(2):E31-6. This process has been improved by decreasing the size of the double-stranded RNA to 20-24 base pairs (to create small-interfering RNAs or siRNAs) that “switched off” genes in mammalian cells without initiating an acute phase response, i.e., a host defense mechanism that often results in cell death (see, e.g., Caplen et al. Proc Natl Acad Sci U S A. 2001 Aug 14;98(17):9742-7 and Elbashir et al. Methods 2002 Feb;26(2):199-213). There is increasing evidence of post-transcriptional gene silencing by RNA interference (RNAi) for inhibiting targeted expression in mammalian cells at the mRNA level, in human cells. There is additional evidence of effective methods for inhibiting the proliferation and migration of tumor cells in human patients, and for inhibiting metastatic cancer development (see, e.g., U.S. Patent Application No. US2001000993183; Caplen et al. Proc Natl Acad Sci U S A; and Abderrahmani et al. Mol Cell Biol 2001 Nov21(21):7256-67).

[0121] An “siRNA” or “RNAi” refers to a nucleic acid that forms a double stranded RNA and has the ability to reduce or inhibit expression of a gene or target gene when the siRNA is delivered to or expressed in the same cell as the gene or target gene. “siRNA” refers to short double-stranded RNA formed by the complementary strands. Complementary portions of the siRNA that hybridize to form the

double stranded molecule often have substantial or complete identity to the target molecule sequence. In one embodiment, an siRNA refers to a nucleic acid that has substantial or complete identity to a target gene and forms a double stranded siRNA.

[0122] When designing the siRNA molecules, the targeted region often is selected from a given DNA sequence beginning 50 to 100 nucleotides downstream of the start codon. See, e.g., Elbashir et al., *Methods* 26:199-213 (2002). Initially, 5' or 3' UTRs and regions nearby the start codon were avoided assuming that UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP or RISC endonuclease complex. Sometimes regions of the target 23 nucleotides in length conforming to the sequence motif AA(N19)TT (N, an nucleotide), and regions with approximately 30% to 70% G/C-content (often about 50% G/C-content) often are selected. If no suitable sequences are found, the search often is extended using the motif NA(N21). The sequence of the sense siRNA sometimes corresponds to (N19) TT or N21 (position 3 to 23 of the 23-nt motif), respectively. In the latter case, the 3' end of the sense siRNA often is converted to TT. The rationale for this sequence conversion is to generate a symmetric duplex with respect to the sequence composition of the sense and antisense 3' overhangs. The antisense siRNA is synthesized as the complement to position 1 to 21 of the 23-nt motif. Because position 1 of the 23-nt motif is not recognized sequence-specifically by the antisense siRNA, the 3'-most nucleotide residue of the antisense siRNA can be chosen deliberately. However, the penultimate nucleotide of the antisense siRNA (complementary to position 2 of the 23-nt motif) often is complementary to the targeted sequence. For simplifying chemical synthesis, TT often is utilized. siRNAs corresponding to the target motif NAR(N17)YNN, where R is purine (A,G) and Y is pyrimidine (C,U), often are selected. Respective 21 nucleotide sense and antisense siRNAs often begin with a purine nucleotide and can also be expressed from pol III expression vectors without a change in targeting site. Expression of RNAs from pol III promoters often is efficient when the first transcribed nucleotide is a purine.

[0123] The sequence of the siRNA can correspond to the full length target gene, or a subsequence thereof. Often, the siRNA is about 15 to about 50 nucleotides in length (e.g., each complementary sequence of the double stranded siRNA is 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, sometimes about 20-30 nucleotides in length or about 20-25 nucleotides in length, e.g., 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length. The siRNA sometimes is about 21 nucleotides in length. Methods of using siRNA are well known in the art, and specific siRNA molecules may be purchased from a number of companies including Dharmacon Research, Inc.

[0124] Antisense, ribozyme, RNAi and siRNA nucleic acids can be altered to form modified nucleic acid molecules. The nucleic acids can be altered at base moieties, sugar moieties or phosphate backbone moieties to improve stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of nucleic acid molecules can be modified to generate peptide nucleic acids (see

Hyrup et al., *Bioorganic & Medicinal Chemistry* 4 (1): 5-23 (1996)). As used herein, the terms “peptide nucleic acid” or “PNA” refers to a nucleic acid mimic such as a DNA mimic, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of a PNA can allow for specific hybridization to DNA and RNA under conditions of low ionic strength. Synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described, for example, in Hyrup et al., (1996) *supra* and Perry-O’Keefe et al., *Proc. Natl. Acad. Sci.* 93: 14670-675 (1996).

[0125] PNA nucleic acids can be used in prognostic, diagnostic, and therapeutic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNA nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-directed PCR clamping); as “artificial restriction enzymes” when used in combination with other enzymes, (e.g., S1 nucleases (Hyrup (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup et al., (1996) *supra*; Perry-O’Keefe *supra*).

[0126] In other embodiments, oligonucleotides may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across cell membranes (see e.g., Letsinger et al., *Proc. Natl. Acad. Sci. USA* 86: 6553-6556 (1989); Lemaitre et al., *Proc. Natl. Acad. Sci. USA* 84: 648-652 (1987); PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, e.g., Krol et al., *Bio-Techniques* 6: 958-976 (1988)) or intercalating agents. (See, e.g., Zon, *Pharm. Res.* 5: 539-549 (1988)). To this end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

[0127] Also included herein are molecular beacon oligonucleotide primer and probe molecules having one or more regions complementary to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence, or a substantially identical sequence thereof, two complementary regions one having a fluorophore and one a quencher such that the molecular beacon is useful for quantifying the presence of the nucleic acid in a sample. Molecular beacon nucleic acids are described, for example, in Lizardi et al., U.S. Patent No. 5,854,033; Nazarenko et al., U.S. Patent No. 5,866,336, and Livak et al., U.S. Patent 5,876,930.

Antibodies

[0128] The term “antibody” as used herein refers to an immunoglobulin molecule or immunologically active portion thereof, i.e., an antigen-binding portion. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab’)₂ fragments which can be

generated by treating the antibody with an enzyme such as pepsin. An antibody sometimes is a polyclonal, monoclonal, recombinant (e.g., a chimeric or humanized), fully human, non-human (e.g., murine), or a single chain antibody. An antibody may have effector function and can fix complement, and is sometimes coupled to a toxin or imaging agent.

[0129] A full-length polypeptide or antigenic peptide fragment encoded by a nucleotide sequence referenced herein can be used as an immunogen or can be used to identify antibodies made with other immunogens, e.g., cells, membrane preparations, and the like. An antigenic peptide often includes at least 8 amino acid residues of the amino acid sequences encoded by a nucleotide sequence referenced herein, or substantially identical sequence thereof, and encompasses an epitope. Antigenic peptides sometimes include 10 or more amino acids, 15 or more amino acids, 20 or more amino acids, or 30 or more amino acids. Hydrophilic and hydrophobic fragments of polypeptides sometimes are used as immunogens.

[0130] Epitopes encompassed by the antigenic peptide are regions located on the surface of the polypeptide (e.g., hydrophilic regions) as well as regions with high antigenicity. For example, an Emini surface probability analysis of the human polypeptide sequence can be used to indicate the regions that have a particularly high probability of being localized to the surface of the polypeptide and are thus likely to constitute surface residues useful for targeting antibody production. The antibody may bind an epitope on any domain or region on polypeptides described herein.

[0131] Also, chimeric, humanized, and completely human antibodies are useful for applications which include repeated administration to subjects. Chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, can be made using standard recombinant DNA techniques. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson et al International Application No. PCT/US86/02269; Akira, et al European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al European Patent Application 173,494; Neuberger et al PCT International Publication No. WO 86/01533; Cabilly et al U.S. Patent No. 4,816,567; Cabilly et al European Patent Application 125,023; Better et al., Science 240: 1041-1043 (1988); Liu et al., Proc. Natl. Acad. Sci. USA 84: 3439-3443 (1987); Liu et al., J. Immunol. 139: 3521-3526 (1987); Sun et al., Proc. Natl. Acad. Sci. USA 84: 214-218 (1987); Nishimura et al., Canc. Res. 47: 999-1005 (1987); Wood et al., Nature 314: 446-449 (1985); and Shaw et al., J. Natl. Cancer Inst. 80: 1553-1559 (1988); Morrison, S. L., Science 229: 1202-1207 (1985); Oi et al., BioTechniques 4: 214 (1986); Winter U.S. Patent 5,225,539; Jones et al., Nature 321: 552-525 (1986); Verhoeyan et al., Science 239: 1534; and Beidler et al., J. Immunol. 141: 4053-4060 (1988).

[0132] Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice that are incapable of expressing

endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. See, for example, Lonberg and Huszar, *Int. Rev. Immunol.* 13: 65-93 (1995); and U.S. Patent Nos. 5,625,126; 5,633,425; 5,569,825; 5,661,016; and 5,545,806. In addition, companies such as Abgenix, Inc. (Fremont, CA) and Medarex, Inc. (Princeton, NJ), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above. Completely human antibodies that recognize a selected epitope also can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody (e.g., a murine antibody) is used to guide the selection of a completely human antibody recognizing the same epitope. This technology is described for example by Jespers et al., *Bio/Technology* 12: 899-903 (1994).

[0133] An antibody can be a single chain antibody. A single chain antibody (scFV) can be engineered (see, e.g., Colcher et al., *Ann. N Y Acad. Sci.* 880: 263-80 (1999); and Reiter, *Clin. Cancer Res.* 2: 245-52 (1996)). Single chain antibodies can be dimerized or multimerized to generate multivalent antibodies having specificities for different epitopes of the same target polypeptide.

[0134] Antibodies also may be selected or modified so that they exhibit reduced or no ability to bind an Fc receptor. For example, an antibody may be an isotype or subtype, fragment or other mutant, which does not support binding to an Fc receptor (e.g., it has a mutagenized or deleted Fc receptor binding region).

[0135] Also, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1 dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thiotepa chlorambucil, melphalan, carmustine (BCNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

[0136] Antibody conjugates can be used for modifying a given biological response. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, gamma-interferon, alpha-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for

example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors. Also, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980, for example.

[0137] An antibody (e.g., monoclonal antibody) can be used to isolate target polypeptides by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, an antibody can be used to detect a target polypeptide (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the polypeptide. Antibodies can be used diagnostically to monitor polypeptide levels in tissue as part of a clinical testing procedure, e.g., to determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance (i.e., antibody labeling). Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aquorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H . Also, an antibody can be utilized as a test molecule for determining whether it can treat osteoarthritis, and as a therapeutic for administration to a subject for treating osteoarthritis.

[0138] An antibody can be made by immunizing with a purified antigen, or a fragment thereof, e.g., a fragment described herein, a membrane associated antigen, tissues, e.g., crude tissue preparations, whole cells, preferably living cells, lysed cells, or cell fractions.

[0139] Included herein are antibodies which bind only a native polypeptide, only denatured or otherwise non-native polypeptide, or which bind both, as well as those having linear or conformational epitopes. Conformational epitopes sometimes can be identified by selecting antibodies that bind to native but not denatured polypeptide. Also featured are antibodies that specifically bind to a polypeptide variant associated with osteoarthritis.

Methods for Identifying Candidate Therapeutics for Treating Osteoarthritis

[0140] Current therapies for the treatment of osteoarthritis have limited efficacy, limited tolerability and significant mechanism-based side effects, and few of the available therapies adequately address underlying defects. Current therapeutic approaches were largely developed in the absence of defined molecular targets or even a solid understanding of disease pathogenesis. Therefore, provided are

methods of identifying candidate therapeutics that target biochemical pathways related to the development of osteoarthritis.

[0141] Thus, featured herein are methods for identifying a candidate therapeutic for treating osteoarthritis. The methods comprise contacting a test molecule with a target molecule in a system. A “target molecule” as used herein refers to a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid, a substantially identical nucleic acid thereof, or a fragment thereof, and an encoded polypeptide of the foregoing. The methods also comprise determining the presence or absence of an interaction between the test molecule and the target molecule, where the presence of an interaction between the test molecule and the nucleic acid or polypeptide identifies the test molecule as a candidate osteoarthritis therapeutic. The interaction between the test molecule and the target molecule may be quantified.

[0142] Test molecules and candidate therapeutics include, but are not limited to, compounds, antisense nucleic acids, siRNA molecules, ribozymes, polypeptides or proteins encoded by a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleotide sequence, or a substantially identical sequence or fragment thereof, and immunotherapeutics (e.g., antibodies and HLA-presented polypeptide fragments). A test molecule or candidate therapeutic may act as a modulator of target molecule concentration or target molecule function in a system. A “modulator” may agonize (i.e., up-regulates) or antagonize (i.e., down-regulates) a target molecule concentration partially or completely in a system by affecting such cellular functions as DNA replication and/or DNA processing (e.g., DNA methylation or DNA repair), RNA transcription and/or RNA processing (e.g., removal of intronic sequences and/or translocation of spliced mRNA from the nucleus), polypeptide production (e.g., translation of the polypeptide from mRNA), and/or polypeptide post-translational modification (e.g., glycosylation, phosphorylation, and proteolysis of pro-polypeptides). A modulator may also agonize or antagonize a biological function of a target molecule partially or completely, where the function may include adopting a certain structural conformation, interacting with one or more binding partners, ligand binding, catalysis (e.g., phosphorylation, dephosphorylation, hydrolysis, methylation, and isomerization), and an effect upon a cellular event (e.g., effecting progression of osteoarthritis).

[0143] As used herein, the term “system” refers to a cell free *in vitro* environment and a cell-based environment such as a collection of cells, a tissue, an organ, or an organism. A system is “contacted” with a test molecule in a variety of manners, including adding molecules in solution and allowing them to interact with one another by diffusion, cell injection, and any administration routes in an animal. As used herein, the term “interaction” refers to an effect of a test molecule on test molecule, where the effect sometimes is binding between the test molecule and the target molecule, and sometimes is an observable change in cells, tissue, or organism.

[0144] There are many standard methods for detecting the presence or absence of interaction between a test molecule and a target molecule. For example, titrametric, acidimetric, radiometric, NMR,

monolayer, polarographic, spectrophotometric, fluorescent, and ESR assays probative of a target molecule interaction may be utilized. Any modulator can be utilized in methods for detecting an interaction. For example, proteasome modulators (e.g., *PSMB1* includes a proteasome domain) are described in WO-2004014882 and Roesel et al. Proceedings of the American Association of Cancer Research 2003, 44:1st Ed (Abs 1769), and bortezomib (Velcade, MLN-341, LDP-341 and PS-341), a ubiquitin proteasome inhibitor, is used for the treatment of multiple myeloma.

[0145] Test molecule/target molecule interactions can be detected and/or quantified using assays known in the art. For example, an interaction can be determined by labeling the test molecule and/or the target molecule, where the label is covalently or non-covalently attached to the test molecule or target molecule. The label is sometimes a radioactive molecule such as ^{125}I , ^{131}I , ^{35}S or ^3H , which can be detected by direct counting of radioemission or by scintillation counting. Also, enzymatic labels such as horseradish peroxidase, alkaline phosphatase, or luciferase may be utilized where the enzymatic label can be detected by determining conversion of an appropriate substrate to product. In addition, presence or absence of an interaction can be determined without labeling. For example, a microphysiometer (e.g., Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indication of an interaction between a test molecule and target molecule (McConnell, H. M. *et al.*, *Science* 257: 1906-1912 (1992)).

[0146] In cell-based systems, cells typically include a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid, an encoded polypeptide, or substantially identical nucleic acid or polypeptide thereof, and are often of mammalian origin, although the cell can be of any origin. Whole cells, cell homogenates, and cell fractions (e.g., cell membrane fractions) can be subjected to analysis. Where interactions between a test molecule with a target polypeptide are monitored, soluble and/or membrane bound forms of the polypeptide may be utilized. Where membrane-bound forms of the polypeptide are used, it may be desirable to utilize a solubilizing agent. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton[®] X-100, Triton[®] X-114, Thesit[®], Isotridecypoly(ethylene glycol ether)_n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.

[0147] An interaction between a test molecule and target molecule also can be detected by monitoring fluorescence energy transfer (FET) (see, e.g., Lakowicz *et al.*, U.S. Patent No. 5,631,169; Stavrianopoulos *et al.* U.S. Patent No. 4,868,103). A fluorophore label on a first, "donor" molecule is selected such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second,

“acceptor” molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the “donor” polypeptide molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the “acceptor” molecule label may be differentiated from that of the “donor”. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, the spatial relationship between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the “acceptor” molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter).

[0148] In another embodiment, determining the presence or absence of an interaction between a test molecule and a target molecule can be effected by monitoring surface plasmon resonance (*see, e.g.*, Sjolander & Urbanicz, *Anal. Chem.* 63: 2338-2345 (1991) and Szabo *et al.*, *Curr. Opin. Struct. Biol.* 5: 699-705 (1995)). “Surface plasmon resonance” or “biomolecular interaction analysis (BIA)” can be utilized to detect biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

[0149] In another embodiment, the target molecule or test molecules are anchored to a solid phase, facilitating the detection of target molecule/test molecule complexes and separation of the complexes from free, uncomplexed molecules. The target molecule or test molecule is immobilized to the solid support. In an embodiment, the target molecule is anchored to a solid surface, and the test molecule, which is not anchored, can be labeled, either directly or indirectly, with detectable labels discussed herein.

[0150] It may be desirable to immobilize a target molecule, an anti-target molecule antibody, and/or test molecules to facilitate separation of target molecule/test molecule complexes from uncomplexed forms, as well as to accommodate automation of the assay. The attachment between a test molecule and/or target molecule and the solid support may be covalent or non-covalent (*see, e.g.*, U.S. Patent No. 6,022,688 for non-covalent attachments). The solid support may be one or more surfaces of the system, such as one or more surfaces in each well of a microtiter plate, a surface of a silicon wafer, a surface of a bead (*see, e.g.*, Lam, *Nature* 354: 82-84 (1991)) that is optionally linked to another solid support, or a channel in a microfluidic device, for example. Types of solid supports, linker molecules for covalent and non-covalent attachments to solid supports, and methods for immobilizing nucleic acids and other molecules to solid supports are well known (*see, e.g.*, U.S. Patent Nos. 6,261,776; 5,900,481; 6,133,436; and 6,022,688; and WIPO publication WO 01/18234).

[0151] In an embodiment, target molecule may be immobilized to surfaces via biotin and streptavidin. For example, biotinylated target polypeptide can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In another embodiment, a target polypeptide can be prepared as a fusion polypeptide. For example, glutathione-S-transferase/target polypeptide fusion can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivitized microtiter plates, which are then combined with a test molecule under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, or the matrix is immobilized in the case of beads, and complex formation is determined directly or indirectly as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of target molecule binding or activity is determined using standard techniques.

[0152] In an embodiment, the non-immobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) under conditions such that a significant percentage of complexes formed will remain immobilized to the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of manners. Where the previously non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface, *e.g.*, by adding a labeled antibody specific for the immobilized component, where the antibody, in turn, can be directly labeled or indirectly labeled with, *e.g.*, a labeled anti-Ig antibody.

[0153] In another embodiment, an assay is performed utilizing antibodies that specifically bind target molecule or test molecule but do not interfere with binding of the target molecule to the test molecule. Such antibodies can be derivitized to a solid support, and unbound target molecule may be immobilized by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

[0154] Cell free assays also can be conducted in a liquid phase. In such an assay, reaction products are separated from unreacted components, by any of a number of standard techniques, including but not limited to: differential centrifugation (*see, e.g.*, Rivas, G., and Minton, *Trends Biochem Sci Aug;18(8):284-7 (1993)*); chromatography (gel filtration chromatography, ion-exchange chromatography); electrophoresis (*see, e.g.*, Ausubel *et al.*, eds. *Current Protocols in Molecular Biology*, J. Wiley: New York (1999)); and immunoprecipitation (*see, e.g.*, Ausubel *et al.*, eds., *supra*). Media and

chromatographic techniques are known to one skilled in the art (*see, e.g., Heegaard, J Mol. Recognit. Winter; 11(1-6): 141-8 (1998); Hage & Tweed, J. Chromatogr. B Biomed. Sci. Appl. Oct 10; 699 (1-2): 499-525 (1997)*). Further, fluorescence energy transfer may also be conveniently utilized, as described herein, to detect binding without further purification of the complex from solution.

[0155] In another embodiment, modulators of target molecule expression are identified. For example, a cell or cell free mixture is contacted with a candidate compound and the expression of target mRNA or target polypeptide is evaluated relative to the level of expression of target mRNA or target polypeptide in the absence of the candidate compound. When expression of target mRNA or target polypeptide is greater in the presence of the candidate compound than in its absence, the candidate compound is identified as an agonist of target mRNA or target polypeptide expression. Alternatively, when expression of target mRNA or target polypeptide is less (*e.g., less with statistical significance*) in the presence of the candidate compound than in its absence, the candidate compound is identified as an antagonist or inhibitor of target mRNA or target polypeptide expression. The level of target mRNA or target polypeptide expression can be determined by methods described herein.

[0156] In another embodiment, binding partners that interact with a target molecule are detected. The target molecules can interact with one or more cellular or extracellular macromolecules, such as polypeptides *in vivo*, and these interacting molecules are referred to herein as “binding partners.” Binding partners can agonize or antagonize target molecule biological activity. Also, test molecules that agonize or antagonize interactions between target molecules and binding partners can be useful as therapeutic molecules as they can up-regulate or down-regulated target molecule activity *in vivo* and thereby treat osteoarthritis.

[0157] Binding partners of target molecules can be identified by methods known in the art. For example, binding partners may be identified by lysing cells and analyzing cell lysates by electrophoretic techniques. Alternatively, a two-hybrid assay or three-hybrid assay can be utilized (*see, e.g., U.S. Patent No. 5,283,317; Zervos et al., Cell 72:223-232 (1993); Madura et al., J. Biol. Chem. 268: 12046-12054 (1993); Bartel et al., Biotechniques 14: 920-924 (1993); Iwabuchi et al., Oncogene 8: 1693-1696 (1993); and Brent WO94/10300*). A two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. The assay often utilizes two different DNA constructs. In one construct, a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid (sometimes referred to as the “bait”) is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g., GAL-4*). In another construct, a DNA sequence from a library of DNA sequences that encodes a potential binding partner (sometimes referred to as the “prey”) is fused to a gene that encodes an activation domain of the known transcription factor. Sometimes, a *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1* or *ERG* nucleic acid can be fused to the activation domain. If the “bait” and the “prey” molecules interact *in vivo*, the DNA-binding and activation domains of the transcription factor are

brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to identify the potential binding partner.

[0158] In an embodiment for identifying test molecules that antagonize or agonize complex formation between target molecules and binding partners, a reaction mixture containing the target molecule and the binding partner is prepared, under conditions and for a time sufficient to allow complex formation. The reaction mixture often is provided in the presence or absence of the test molecule. The test molecule can be included initially in the reaction mixture, or can be added at a time subsequent to the addition of the target molecule and its binding partner. Control reaction mixtures are incubated without the test molecule or with a placebo. Formation of any complexes between the target molecule and the binding partner then is detected. Decreased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule antagonizes target molecule/binding partner complex formation. Alternatively, increased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule agonizes target molecule/binding partner complex formation. In another embodiment, complex formation of target molecule/binding partner can be compared to complex formation of mutant target molecule/binding partner (*e.g.*, amino acid modifications in a target polypeptide). Such a comparison can be important in those cases where it is desirable to identify test molecules that modulate interactions of mutant but not non-mutated target gene products.

[0159] The assays can be conducted in a heterogeneous or homogeneous format. In heterogeneous assays, target molecule and/or the binding partner are immobilized to a solid phase, and complexes are detected on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the molecules being tested. For example, test compounds that agonize target molecule/binding partner interactions can be identified by conducting the reaction in the presence of the test molecule in a competition format. Alternatively, test molecules that agonize preformed complexes, *e.g.*, molecules with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed.

[0160] In a heterogeneous assay embodiment, the target molecule or the binding partner is anchored onto a solid surface (*e.g.*, a microtiter plate), while the non-anchored species is labeled, either directly or indirectly. The anchored molecule can be immobilized by non-covalent or covalent attachments. Alternatively, an immobilized antibody specific for the molecule to be anchored can be used to anchor the molecule to the solid surface. The partner of the immobilized species is exposed to the coated surface with or without the test molecule. After the reaction is complete, unreacted components are removed

(*e.g.*, by washing) such that a significant portion of any complexes formed will remain immobilized on the solid surface. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface is indicative of complex. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored to the surface; *e.g.*, by using a labeled antibody specific for the initially non-immobilized species. Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

[0161] In another embodiment, the reaction can be conducted in a liquid phase in the presence or absence of test molecule, where the reaction products are separated from unreacted components, and the complexes are detected (*e.g.*, using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes). Again, depending upon the order of addition of reactants to the liquid phase, test compounds that inhibit complex or that disrupt preformed complexes can be identified.

[0162] In an alternate embodiment, a homogeneous assay can be utilized. For example, a preformed complex of the target gene product and the interactive cellular or extracellular binding partner product is prepared. One or both of the target molecule or binding partner is labeled, and the signal generated by the label(s) is quenched upon complex formation (*e.g.*, U.S. Patent No. 4,109,496 that utilizes this approach for immunoassays). Addition of a test molecule that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances that disrupt target molecule/binding partner complexes can be identified.

[0163] Candidate therapeutics for treating osteoarthritis are identified from a group of test molecules that interact with a target molecule. Test molecules are normally ranked according to the degree with which they modulate (*e.g.*, agonize or antagonize) a function associated with the target molecule (*e.g.*, DNA replication and/or processing, RNA transcription and/or processing, polypeptide production and/or processing, and/or biological function/activity), and then top ranking modulators are selected. Also, pharmacogenomic information described herein can determine the rank of a modulator. The top 10% of ranked test molecules often are selected for further testing as candidate therapeutics, and sometimes the top 15%, 20%, or 25% of ranked test molecules are selected for further testing as candidate therapeutics. Candidate therapeutics typically are formulated for administration to a subject.

Therapeutic Formulations

[0164] Formulations and pharmaceutical compositions typically include in combination with a pharmaceutically acceptable carrier one or more target molecule modulators. The modulator often is a test molecule identified as having an interaction with a target molecule by a screening method described above. The modulator may be a compound, an antisense nucleic acid, a ribozyme, an antibody, or a

binding partner. Also, formulations may comprise a target polypeptide or fragment thereof in combination with a pharmaceutically acceptable carrier.

[0165] As used herein, the term “pharmaceutically acceptable carrier” includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions. Pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

[0166] A pharmaceutical composition typically is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerin, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

[0167] Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules, *e.g.*, gelatin capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

[0168] Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can

be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

[0169] Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0170] For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

[0171] Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art. Molecules can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

[0172] In one embodiment, active molecules are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. Materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to

viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

[0173] It is advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier.

[0174] Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀. Molecules which exhibit high therapeutic indices are preferred. While molecules that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[0175] The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such molecules lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any molecules used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC₅₀ (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

[0176] As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, sometimes about 0.01 to 25 mg/kg body weight, often about 0.1 to 20 mg/kg body weight, and more often about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The protein or polypeptide can be administered one time per week for between about 1 to 10 weeks, sometimes between 2 to 8 weeks, often between about 3 to 7 weeks, and more often for about 4, 5, or 6 weeks. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a

therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

[0177] With regard to polypeptide formulations, featured herein is a method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject with a first polypeptide, where the subject comprises a second polypeptide having one or more polymorphic variations associated with cancer, and where the first polypeptide comprises fewer polymorphic variations associated with cancer than the second polypeptide. The first and second polypeptides are encoded by a nucleic acid which comprises a nucleotide sequence in SEQ ID NO: 1-12; a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence referenced in SEQ ID NO: 1-12; a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-12 and a nucleotide sequence 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-12. The subject often is a human.

[0178] For antibodies, a dosage of 0.1 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg) is often utilized. If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is often appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (*e.g.*, into the brain). A method for lipidation of antibodies is described by Cruikshank *et al.*, *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193 (1997).

[0179] Antibody conjugates can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors. Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

[0180] For compounds, exemplary doses include milligram or microgram amounts of the compound per kilogram of subject or sample weight, for example, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is understood that appropriate

doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid described herein, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

[0181] With regard to nucleic acid formulations, gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see, e.g.*, U.S. Patent 5,328,470) or by stereotactic injection (*see e.g.*, Chen *et al.*, (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). Pharmaceutical preparations of gene therapy vectors can include a gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells (*e.g.*, retroviral vectors) the pharmaceutical preparation can include one or more cells which produce the gene delivery system. Examples of gene delivery vectors are described herein.

Therapeutic Methods

[0182] A therapeutic formulation described above can be administered to a subject in need of a therapeutic for inducing a desired biological response. Therapeutic formulations can be administered by any of the paths described herein. With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from pharmacogenomic analyses described herein.

[0183] As used herein, the term “treatment” is defined as the application or administration of a therapeutic formulation to a subject, or application or administration of a therapeutic agent to an isolated tissue or cell line from a subject with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect osteoarthritis, symptoms of osteoarthritis or a predisposition towards osteoarthritis. A therapeutic formulation includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides. Administration of a therapeutic formulation can occur prior to the manifestation of symptoms characteristic of osteoarthritis, such that osteoarthritis is prevented or delayed in its progression. The appropriate therapeutic composition can be determined based on screening assays described herein.

[0184] As discussed, successful treatment of osteoarthritis can be brought about by techniques that serve to agonize target molecule expression or function, or alternatively, antagonize target molecule

expression or function. These techniques include administration of modulators that include, but are not limited to, small organic or inorganic molecules; antibodies (including, for example, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and Fab, F(ab')₂ and Fab expression library fragments, scFV molecules, and epitope-binding fragments thereof); and peptides, phosphopeptides, or polypeptides.

[0185] Further, antisense and ribozyme molecules that inhibit expression of the target gene can also be used to reduce the level of target gene expression, thus effectively reducing the level of target gene activity. Still further, triple helix molecules can be utilized in reducing the level of target gene activity. Antisense, ribozyme and triple helix molecules are discussed above. It is possible that the use of antisense, ribozyme, and/or triple helix molecules to reduce or inhibit mutant gene expression can also reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles, such that the concentration of normal target gene product present can be lower than is necessary for a normal phenotype. In such cases, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity can be introduced into cells via gene therapy method. Alternatively, in instances in that the target gene encodes an extracellular polypeptide, it can be preferable to co-administer normal target gene polypeptide into the cell or tissue in order to maintain the requisite level of cellular or tissue target gene activity.

[0186] Another method by which nucleic acid molecules may be utilized in treating or preventing osteoarthritis is use of aptamer molecules specific for target molecules. Aptamers are nucleic acid molecules having a tertiary structure which permits them to specifically bind to ligands (*see, e.g., Osborne, et al., Curr. Opin. Chem. Biol.* 1(1): 5-9 (1997); and Patel, D. J., *Curr. Opin. Chem. Biol. Jun;1(1): 32-46 (1997)*).

[0187] Yet another method of utilizing nucleic acid molecules for osteoarthritis treatment is gene therapy, which can also be referred to as allele therapy. Provided herein is a gene therapy method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject or from the subject with a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-12). Genomic DNA in the subject comprises a second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-4). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human. Allele therapy methods often are utilized in conjunction with a

method of first determining whether a subject has genomic DNA that includes polymorphic variants associated with osteoarthritis.

[0188] In another allele therapy embodiment, provided herein is a method which comprises contacting one or more cells in the subject or from the subject with a polypeptide encoded by a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to the nucleotide sequence of SEQ ID NO: 1-12). Genomic DNA in the subject comprises a second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-4). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human.

[0189] For antibody-based therapies, antibodies can be generated that are both specific for target molecules and that reduce target molecule activity. Such antibodies may be administered in instances where antagonizing a target molecule function is appropriate for the treatment of osteoarthritis.

[0190] In circumstances where stimulating antibody production in an animal or a human subject by injection with a target molecule is harmful to the subject, it is possible to generate an immune response against the target molecule by use of anti-idiotypic antibodies (*see, e.g., Herlyn, Ann. Med.; 31(1): 66-78 (1999); and Bhattacharya-Chatterjee & Foon, Cancer Treat. Res.; 94: 51-68 (1998)*). Introducing an anti-idiotypic antibody to a mammal or human subject often stimulates production of anti-anti-idiotypic antibodies, which typically are specific to the target molecule. Vaccines directed to osteoarthritis also may be generated in this fashion.

[0191] In instances where the target molecule is intracellular and whole antibodies are used, internalizing antibodies may be preferred. Lipofectin or liposomes can be used to deliver the antibody or a fragment of the Fab region that binds to the target antigen into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target antigen is preferred. For example, peptides having an amino acid sequence corresponding to the Fv region of the antibody can be used. Alternatively, single chain neutralizing antibodies that bind to intracellular target antigens can also be administered. Such single chain antibodies can be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population (*see, e.g., Marasco et al., Proc. Natl. Acad. Sci. USA 90: 7889-7893 (1993)*).

[0192] Modulators can be administered to a patient at therapeutically effective doses to treat osteoarthritis. A therapeutically effective dose refers to an amount of the modulator sufficient to result in amelioration of symptoms of osteoarthritis. Toxicity and therapeutic efficacy of modulators can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g., for*

determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀. Modulators that exhibit large therapeutic indices are preferred. While modulators that exhibit toxic side effects can be used, care should be taken to design a delivery system that targets such molecules to the site of affected tissue in order to minimize potential damage to uninfected cells, thereby reducing side effects.

[0193] Data obtained from cell culture assays and animal studies can be used in formulating a range of dosages for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC₅₀ (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma can be measured, for example, by high performance liquid chromatography.

[0194] Another example of effective dose determination for an individual is the ability to directly assay levels of “free” and “bound” compound in the serum of the test subject. Such assays may utilize antibody mimics and/or “biosensors” that have been created through molecular imprinting techniques. Molecules that modulate target molecule activity are used as a template, or “imprinting molecule”, to spatially organize polymerizable monomers prior to their polymerization with catalytic reagents. The subsequent removal of the imprinted molecule leaves a polymer matrix which contains a repeated “negative image” of the compound and is able to selectively rebind the molecule under biological assay conditions. A detailed review of this technique can be seen in Ansell *et al.*, *Current Opinion in Biotechnology* 7: 89-94 (1996) and in Shea, *Trends in Polymer Science* 2: 166-173 (1994). Such “imprinted” affinity matrixes are amenable to ligand-binding assays, whereby the immobilized monoclonal antibody component is replaced by an appropriately imprinted matrix. An example of the use of such matrixes in this way can be seen in Vlatakis, *et al.*, *Nature* 361: 645-647 (1993). Through the use of isotope-labeling, the “free” concentration of compound which modulates target molecule expression or activity readily can be monitored and used in calculations of IC₅₀. Such “imprinted” affinity matrixes can also be designed to include fluorescent groups whose photon-emitting properties measurably change upon local and selective binding of target compound. These changes readily can be assayed in real time using appropriate fiberoptic devices, in turn allowing the dose in a test subject to be quickly optimized based on its individual IC₅₀. An example of such a “biosensor” is discussed in Kriz *et al.*, *Analytical Chemistry* 67: 2142-2144 (1995).

[0195] The examples set forth below are intended to illustrate but not limit the invention.

Examples

[0196] In the following studies a group of subjects was selected according to specific parameters relating to osteoarthritis. Nucleic acid samples obtained from individuals in the study group were subjected to genetic analysis, which identified associations between osteoarthritis and polymorphisms in the following genes or regions: *chromosome 6 (6q27)*, *ELP3*, *CHDC1*, and *ERG* (herein referred to as “targets”). The polymorphisms were genotyped again in two replication cohorts consisting of individuals selected for OA. In addition, SNPs proximal to the incident polymorphisms were identified and allelotyped in OA case and control pools. Methods are described for producing *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1*, and *ERG* polypeptide and polypeptide variants thereof *in vitro* or *in vivo*; *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1*, and *ERG* nucleic acids or polypeptides and variants thereof are utilized for screening test molecules for those that interact with *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1*, and *ERG* molecules. Test molecules identified as interactors with *PSMB1*, *TBP*, *PDCD2*, *ELP3*, *CHDC1*, and *ERG* molecules and variants thereof are further screened *in vivo* to determine whether they treat osteoarthritis.

Example 1

Samples and Pooling Strategies

Sample Selection

[0197] Blood samples were collected from individuals diagnosed with knee osteoarthritis, which were referred to as case samples. Also, blood samples were collected from individuals not diagnosed with knee osteoarthritis as gender and age-matched controls. A database was created that listed all phenotypic trait information gathered from individuals for each case and control sample. Genomic DNA was extracted from each of the blood samples for genetic analyses.

DNA Extraction from Blood Samples

[0198] Six to ten milliliters of whole blood was transferred to a 50 ml tube containing 27 ml of red cell lysis solution (RCL). The tube was inverted until the contents were mixed. Each tube was incubated for 10 minutes at room temperature and inverted once during the incubation. The tubes were then centrifuged for 20 minutes at 3000 x g and the supernatant was carefully poured off. 100-200 µl of residual liquid was left in the tube and was pipetted repeatedly to resuspend the pellet in the residual supernatant. White cell lysis solution (WCL) was added to the tube and pipetted repeatedly until completely mixed. While no incubation was normally required, the solution was incubated at 37°C or

room temperature if cell clumps were visible after mixing until the solution was homogeneous. 2 ml of protein precipitation was added to the cell lysate. The mixtures were vortexed vigorously at high speed for 20 sec to mix the protein precipitation solution uniformly with the cell lysate, and then centrifuged for 10 minutes at 3000 x g. The supernatant containing the DNA was then poured into a clean 15 ml tube, which contained 7 ml of 100% isopropanol. The samples were mixed by inverting the tubes gently until white threads of DNA were visible. Samples were centrifuged for 3 minutes at 2000 x g and the DNA was visible as a small white pellet. The supernatant was decanted and 5 ml of 70% ethanol was added to each tube. Each tube was inverted several times to wash the DNA pellet, and then centrifuged for 1 minute at 2000 x g. The ethanol was decanted and each tube was drained on clean absorbent paper. The DNA was dried in the tube by inversion for 10 minutes, and then 1000 µl of 1X TE was added. The size of each sample was estimated, and less TE buffer was added during the following DNA hydration step if the sample was smaller. The DNA was allowed to rehydrate overnight at room temperature, and DNA samples were stored at 2-8°C.

[0199] DNA was quantified by placing samples on a hematology mixer for at least 1 hour. DNA was serially diluted (typically 1:80, 1:160, 1:320, and 1:640 dilutions) so that it would be within the measurable range of standards. 125 µl of diluted DNA was transferred to a clear U-bottom microtitre plate, and 125 µl of 1X TE buffer was transferred into each well using a multichannel pipette. The DNA and 1X TE were mixed by repeated pipetting at least 15 times, and then the plates were sealed. 50 µl of diluted DNA was added to wells A5-H12 of a black flat bottom microtitre plate. Standards were inverted six times to mix them, and then 50 µl of 1X TE buffer was pipetted into well A1, 1000 ng/ml of standard was pipetted into well A2, 500 ng/ml of standard was pipetted into well A3, and 250 ng/ml of standard was pipetted into well A4. PicoGreen (Molecular Probes, Eugene, Oregon) was thawed and freshly diluted 1:200 according to the number of plates that were being measured. PicoGreen was vortexed and then 50µl was pipetted into all wells of the black plate with the diluted DNA. DNA and PicoGreen were mixed by pipetting repeatedly at least 10 times with the multichannel pipette. The plate was placed into a Fluoroskan Ascent Machine (microplate fluorometer produced by Labsystems) and the samples were allowed to incubate for 3 minutes before the machine was run using filter pairs 485 nm excitation and 538 nm emission wavelengths. Samples having measured DNA concentrations of greater than 450 ng/µl were re-measured for conformation. Samples having measured DNA concentrations of 20 ng/µl or less were re-measured for confirmation.

Pooling Strategies – Discovery Cohort

[0200] Samples were derived from the Nottingham knee OA family study (UK) where index cases were identified through a knee replacement registry. Siblings were approached and assessed with knee x-

rays and assigned status as affected or unaffected. In all 1,157 individuals were available. In order to create same-sex pools of appropriate sizes, 335 unrelated female individuals with OA from the Nottingham OA sample were selected for the case pool. The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age. The female case samples and female control samples are described further in Table 1 below.

[0201] A select set of samples from each group were utilized to generate pools, and one pool was created for each group. Each individual sample in a pool was represented by an equal amount of genomic DNA. For example, where 25 ng of genomic DNA was utilized in each PCR reaction and there were 200 individuals in each pool, each individual would provide 125 pg of genomic DNA. Inclusion or exclusion of samples for a pool was based upon the following criteria: the sample was derived from an individual characterized as Caucasian; the sample was derived from an individual of British paternal and maternal descent; case samples were derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic. Control samples were derived from individuals free of OA, family history of OA, and rheumatoid arthritis. Also, sufficient genomic DNA was extracted from each blood sample for all allelotyping and genotyping reactions performed during the study. Phenotype information from each individual was collected and included age of the individual, gender, family history of OA, general medical information (e.g., height, weight, thyroid disease, diabetes, psoriasis, hysterectomy), joint history (previous and current symptoms, joint-related operations, age at onset of symptoms, date of primary diagnosis, age of individual as of primary diagnosis and order of involvement), and knee-related findings (crepitus, restricted passive movement, bony swelling/deformity). Additional knee information included knee history, current symptoms, any major knee injury, menisectomy, knee replacement surgery, age of surgery, and treatment history (including hormone replace therapy (HRT)). Samples that met these criteria were added to appropriate pools based on disease status.

[0202] The selection process yielded the pools set forth in Table 1, which were used in the studies that follow:

TABLE 1

	Female case	Female control
Pool size (Number)	335	335
Pool Criteria (ex: case/control)	control	case

Mean Age (ex: years)	57.21	69.95
-------------------------	-------	-------

Example 2

Association of Polymorphic Variants with Osteoarthritis

[0203] A whole-genome screen was performed to identify particular SNPs associated with occurrence of osteoarthritis. As described in Example 1, two sets of samples were utilized, which included samples from female individuals having knee osteoarthritis (osteoarthritis cases), and samples from female individuals not having knee osteoarthritis (female controls). The initial screen of each pool was performed in an allelotyping study, in which certain samples in each group were pooled. By pooling DNA from each group, an allele frequency for each SNP in each group was calculated. These allele frequencies were then compared to one another. Particular SNPs were considered as being associated with osteoarthritis when allele frequency differences calculated between case and control pools were statistically significant. SNP disease association results obtained from the allelotyping study were then validated by genotyping each associated SNP across all samples from each pool. The results of the genotyping then were analyzed, allele frequencies for each group were calculated from the individual genotyping results, and a p-value was calculated to determine whether the case and control groups had statistically significant differences in allele frequencies for a particular SNP. When the genotyping results agreed with the original allelotyping results, the SNP disease association was considered validated at the genetic level.

SNP Panel Used for Genetic Analyses

[0204] A whole-genome SNP screen began with an initial screen of approximately 25,000 SNPs over each set of disease and control samples using a pooling approach. The pools studied in the screen are described in Example 1. The SNPs analyzed in this study were part of a set of 25,488 SNPs confirmed as being statistically polymorphic as each is characterized as having a minor allele frequency of greater than 10%. The SNPs in the set reside in genes or in close proximity to genes, and many reside in gene exons. Specifically, SNPs in the set are located in exons, introns, and within 5,000 base-pairs upstream of a transcription start site of a gene. In addition, SNPs were selected according to the following criteria: they are located in ESTs; they are located in Locuslink or Ensembl genes; and they are located in Genomatix promoter predictions. SNPs in the set were also selected on the basis of even spacing across the genome, as depicted in Table 2.

[0205] A case-control study design using a whole genome association strategy involving approximately 28,000 single nucleotide polymorphisms (SNPs) was employed. Approximately 25,000

SNPs were evenly spaced in gene-based regions of the human genome with a median inter-marker distance of about 40,000 base pairs. Additionally, approximately 3,000 SNPs causing amino acid substitutions in genes described in the literature as candidates for various diseases were used. The case-control study samples were of female Caucasian origin (British paternal and maternal descent) 670 individuals were equally distributed in two groups: female controls and female cases. The whole genome association approach was first conducted on 2 DNA pools representing the 2 groups. Significant markers were confirmed by individual genotyping.

TABLE 2

<u>General Statistics</u>		<u>Spacing Statistics</u>	
Total # of SNPs	25,488	Median	37,058 bp
# of Exonic SNPs	>4,335 (17%)	Minimum*	1,000 bp
# SNPs with refSNP ID	20,776 (81%)	Maximum*	3,000,000 bp
Gene Coverage	>10,000	Mean	122,412 bp
Chromosome Coverage	All	Std Deviation	373,325 bp
		<i>*Excludes outliers</i>	

Allelotyping and Genotyping Results

[0206] The genetic studies summarized above and described in more detail below identified allelic variants in the target genes that are associated with osteoarthritis.

Assay for Verifying, Allelotyping, and Genotyping SNPs

[0207] A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hME™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0208] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which were used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could manually design primers based on his or her knowledge of the relevant factors and considerations in designing such primers. Table 3 shows PCR primers and Table 4 shows extension primers used for analyzing polymorphisms. The initial PCR amplification reaction was performed in a 5 µl total volume containing 1X PCR buffer with 1.5 mM MgCl₂ (Qiagen), 200 µM each of dATP, dGTP, dCTP, dTTP

(Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

TABLE 3: PCR Primers

SNP Reference	Forward PCR primer	Reverse PCR primer
rs756519	ACGTTGGATGTCTAGAGACACCTGAGGTTG	ACGTTGGATGTGTTTCACTTCAGAGCCCTG
rs1042327	ACGTTGGATGAACTTCACATCACAGCTCCC	ACGTTGGATGCAGAAGTTGGGTTTTCCAGC
rs8770	ACGTTGGATGCTGTCACTGGACACTTTTG	ACGTTGGATGAAAATAGAGGTGCAGAGATG
rs1563055	ACGTTGGATGAGTTCTTTCTCCTCACATTG	ACGTTGGATGCCCTTTAGAAGCACATACTC
rs912428	ACGTTGGATGACTACATCCATTCCAGGGAG	ACGTTGGATGTCAGATCAGAGTGAGTTTAG
rs1888475	ACGTTGGATGACCCCTGGCAAGTGAATTAC	ACGTTGGATGGGGAGGTGGATGTTCTTATC

[0209] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 µl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 µl) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0210] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 4, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

TABLE 4: Extension Primers

SNP Reference	Extend Probe	Termination Mix
rs756519	CAGAGCCCTGTTCTTTGATT	ACG
rs1042327	CATCACAGCTCCCCACCAT	ACT
rs8770	TAGACACTGTGTAAGCAATC	ACG
rs1563055	TTCTCCTCACATTGTTTCTACT	ACG
rs912428	CCATTCCAGGGAGACTCCCA	ACT
rs1888475	GACATCAAATGATTCCCCTGT	ACT

[0211] The MassEXTEND™ reaction was performed in a total volume of 9 µl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was

complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0212] Following incubation, samples were desalted by adding 16 µl of water (total reaction volume was 25 µl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

Genetic Analysis

[0213] Minor allelic frequencies for the polymorphisms set forth in Table A were verified as being 10% or greater using the extension assay described above in a group of samples isolated from 92 individuals originating from the state of Utah in the United States, Venezuela and France (Coriell cell repositories).

[0214] Genotyping results are shown for female pools in Table 5. In Table 5, “AF” refers to allelic frequency; and “F case” and “F control” refer to female case and female control groups, respectively.

TABLE 5: Genotyping Results

SNP Reference	AF F case	AF F control	p-value
rs756519	C = 0.581 T = 0.419	C = 0.656 T = 0.344	0.0055
rs1042327	T = 0.472 C = 0.528	T = 0.563 C = 0.437	0.0012
rs8770	C = 0.529 T = 0.471	C = 0.432 T = 0.568	0.0001
rs1563055	C = 0.653 T = 0.347	C = 0.736 T = 0.264	0.0013
rs912428	T = 0.228 C = 0.772	T = 0.170 C = 0.830	0.0076
rs1888475	A = 0.188 G = 0.812	A = 0.135 G = 0.865	0.0087

[0215] All of the single marker alleles set forth in Table A were considered validated, since the genotyping data agreed with the allelotyping data and each SNP significantly associated with osteoarthritis. Particularly significant associations with osteoarthritis are indicated by a calculated p-value of less than 0.05 for genotype results.

Example 3

Association of Polymorphic Variants with Osteoarthritis in Replication Cohorts

[0216] The single marker polymorphisms set forth in Table A were genotyped again in two replication cohorts consisting of individuals selected for OA.

Sample Selection and Pooling Strategies – Replication Sample 1

[0217] A second case control sample (replication sample #1) was created by using 100 Caucasian female cases from Chingford, UK, and 148 unrelated female cases from the St. Thomas twin study. Cases were defined as having Kellgren-Lawrence (KL) scores of at least 2 in at least one knee x-ray. In addition, 199 male knee replacement cases from Nottingham were included. (For a cohort description, see the Nottingham description provided in Example 1). The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age. The replication sample 1 cohort was used to replicate the initial results. Table 6 below summarizes the selected phenotype data collected from the case and control individuals.

TABLE 6

Phenotype	Female cases (n=248): median (range)/ (n,%)	Male cases (n=199): median (range)/ (n,%)	Female controls (n=313): mean (range)/ (n,%)
Age	59 (39- 73)	66 (45- 73)	55 (50- 72)
Height (cm)	162 (141- 178)	175 (152- 198)	162 (141- 176)
Weight (kg)	68 (51- 123)	86 (62- 127)	64 (40- 111)
Body mass index (kg/m ²)	26 (18- 44)	29 (21- 41)	24 (18- 46)
Kellgren- Lawrence* left knee	0 (63, 26%), 1 (20, 8%), 2 (105, 43%), 3 (58, 23%), 4 (1, 0%)	NA	NA
Kellgren- Lawrence* right knee	0 (43, 7%), 1 (18, 7%), 2 (127, 52%), 3 (57, 23%), 4 (1, 0%)	NA	NA
KL* >2 both knees	No (145, 59%), Yes (101, 41%)	NA	NA
KL* >2 either knee	No (0, 0%), Yes (248, 100%)	NA	NA

* 0: normal, 1: doubtful, 2: definite osteophyte (bony protuberance), 3: joint space narrowing (with or without osteophyte), 4: joint deformity

Sample Selection and Pooling Strategies – Replication Sample 2

[0218] A third case control sample (replication sample #2) was created by using individuals with symptoms of OA from Newfoundland, Canada. These individuals were recruited and examined by rheumatologists. Affected joints were x-rayed and a final diagnosis of definite or probable OA was made according to American College of Rheumatology criteria by a single rheumatologist to avoid any inter-examiner diagnosis variability. Controls were recruited from volunteers without any symptoms from the musculoskeletal system based on a normal joint exam performed by a rheumatologist. Only cases with a diagnosis of definite OA were included in the study. Only individuals of Caucasian origin were included. The cases consisted of 228 individuals with definite knee OA, 106 individuals with definite hip OA, and 74 individuals with hip OA.

TABLE 7

Phenotype	Case	Control
Age at Visit	62.7	52.5
Sex (Female/Male)	227/119	174/101
Knee OA Xray: No	35% (120)	80% (16)
Unknown	1% (4)	0% (0)
Yes	64% (221)	20% (4)
Hip OA Xray: No	63% (215)	80% (16)
Unknown	2% (7)	0% (0)
Yes	35% (121)	20% (4)

Assay for Verifying, Allelotyping, and Genotyping SNPs

[0219] Genotyping of the replication cohorts described in Tables 6 and 7 was performed using the same methods used for the original genotyping, as described herein. A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hME™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0220] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which were used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could manually design primers based on his or her knowledge of the relevant factors and considerations in

designing such primers. Table 3 shows PCR primers and Table 4 shows extension probes used for analyzing (*e.g.*, genotyping) polymorphisms in the replication cohorts. The initial PCR amplification reaction was performed in a 5 μ l total volume containing 1X PCR buffer with 1.5 mM MgCl₂ (Qiagen), 200 μ M each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

[0221] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 μ l volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 μ l) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0222] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 7, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

[0223] The MassEXTEND™ reaction was performed in a total volume of 9 μ l, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0224] Following incubation, samples were desalted by adding 16 μ l of water (total reaction volume was 25 μ l), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and

SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

Genetic Analysis

[0225] Genotyping results for replication cohorts #1 and #2 are provided in Tables 8 and 9, respectively.

TABLE 8

rsID	Replication #1 (Mixed Male/Female cases and Female controls)				Meta-analysis Disc. + Rep #1
	AF OA Con	AF OA Cas	Delta	P-value	P-value
rs756519	0.4	0.43	-0.04	0.140	0.0098
rs1042327	0.49	0.52	-0.03	0.234	0.0430
rs8770	0.51	0.48	0.03	0.303	0.0480
rs1563055	0.31	0.35	-0.04	0.083	0.0002
rs912428	0.86	0.8	0.06	0.004	~0.00001
rs1888475	0.86	0.81	0.04	0.032	0.0002

TABLE 9

rsID	Replication #2 (Newfoundland) (Male/Female cases and controls)				Meta-analysis Disc. + Rep #2
	AF OA Con	AF OA Cas	Delta	P-value	Not Done
rs756519	0.39	0.40	-0.007	0.816	
rs1042327	0.49	0.51	-0.024	0.405	
rs8770	0.53	0.49	0.039	0.195	
rs1563055	0.34	0.34	-0.005	0.864	
rs912428	0.82	0.76	0.058	0.016	
rs1888475	0.80	0.82	-0.025	0.280	

[0226] To combine the evidence for association from multiple sample collections, a meta-analysis procedure was employed. The allele frequencies were compared between cases and controls within the discovery sample, as well as within the replication cohort #1 using the DerSimonian-Laird approach (DerSimonian, R. and N. Laird. 1986. Meta-analysis in clinical trials. Control Clin Trials 7: 177-188.)

[0227] The absence of a statistically significant association in one or more of the replication cohorts should not be interpreted as minimizing the value of the original finding. There are many reasons why a biologically derived association identified in a sample from one population would not replicate in a sample from another population. The most important reason is differences in population history. Due to bottlenecks and founder effects, there may be common disease predisposing alleles present in one population that are relatively rare in another, leading to a lack of association in the candidate region. Also, because common diseases such as arthritis-related disorders are the result of susceptibilities in

many genes and many environmental risk factors, differences in population-specific genetic and environmental backgrounds could mask the effects of a biologically relevant allele. For these and other reasons, statistically strong results in the original, discovery sample that did not replicate in one or more of the replication samples may be further evaluated in additional replication cohorts and experimental systems.

Example 4

Chromosome 6 Region Proximal SNPs

[0228] It has been discovered that SNPs rs756519, rs1042327 and rs8770 on chromosome 6 (6q27) are associated with occurrence of osteoarthritis in subjects. This region contains genes that encode proteasome (prosome, macropain) subunit, beta type, 1 (*PSMB1*), TATA box binding protein (*TBP*), and programmed cell death 2 (*PDCD2*).

[0229] One hundred-nine additional allelic variants proximal to rs756519, rs1042327 and rs8770 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 10. The chromosome positions provided in column four of Table 10 are based on Genome “Build 34” of NCBI’s GenBank.

TABLE 10

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs1474555	6	229	170689279	c/t
rs1474554	6	6310	170695360	a/g
rs10334	6	11840	170700890	g/t
rs10541	6	11870	170700920	a/t
rs3823299	6	12064	170701114	a/g
Rs742348	6	13392	170702442	c/g
rs1474644	6	16354	170705404	a/g
rs1474643	6	16559	170705609	c/t
rs2056970	6	16935	170705985	a/g
rs2223474	6	17616	170706666	c/t
rs2206284	6	17737	170706787	c/t
Rs756519	6	18321	170707371	c/t
Rs756518	6	18453	170707503	a/g
Rs756517	6	18811	170707861	c/t
rs1474642	6	20020	170709070	c/t
rs2038093	6	21662	170710712	c/g
rs2038092	6	23197	170712247	c/g
rs2223473	6	23446	170712496	g/t
Rs760909	6	24339	170713389	g/t
rs2076319	6	25504	170714554	a/g
rs3778589	6	27174	170716224	a/g
rs3800236	6	28008	170717058	a/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs2206286	6	29294	170718344	c/t
rs12717	6	29759	170718809	c/g
rs2179373	6	30832	170719882	a/g
rs3800235	6	44512	170733562	a/c
rs3823298	6	44850	170733900	c/g
rs2076318	6	45884	170734934	a/g
rs2235506	6	46345	170735395	c/t
rs2072916	6	48589	170737639	a/g
rs3734763	6	53371	170742421	a/g
rs3177571	6	53911	170742961	g/t
rs8770	6	53990	170743040	a/g
rs3173219	6	55152	170744202	c/g
Rs960744	6	55667	170744717	c/t
rs2066954	6	58952	170748002	a/c
rs2072917	6	59315	170748365	g/t
rs3173220	6	60029	170749079	a/g
Rs734249	6	61477	170750527	a/c
rs2092310	6	62988	170752038	c/t
rs2092309	6	63090	170752140	c/g
rs1016536	6	64021	170753071	a/c
rs2235506	6	65685	170754735	c/t
rs2076998	6	70220	170759270	a/g
rs2076997	6	70323	170759373	a/c
rs2345478	6	70959	170760009	a/c
rs2021899	6	73436	170762486	c/g
rs2021898	6	82945	170771995	a/g
rs2345682	6	82958	170772008	g/t
rs2345683	6	82961	170772011	c/g
rs2881195	6	82964	170772014	c/t
rs2345684	6	82965	170772015	g/t
rs3046261	6	83006	170772056	-/cttt
rs4083413	6	83025	170772075	c/t
rs4083412	6	83034	170772084	a/g
rs2345685	6	83074	170772124	g/t
rs2021897	6	83132	170772182	g/t
rs4036211	6	83155	170772205	c/t
rs4036212	6	83172	170772222	a/t
rs4036213	6	83174	170772224	g/t
rs2345686	6	83206	170772256	c/t
rs4036214	6	83216	170772266	g/t
rs4036215	6	83234	170772284	g/t
rs2345687	6	83252	170772302	a/g
rs2345688	6	83260	170772310	a/c
rs2881196	6	83263	170772313	a/c
rs3046288	6	83296	170772346	-/at
rs4036216	6	83319	170772369	a/g
rs4036205	6	83322	170772372	c/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs2092307	6	83324	170772374	a/c
rs4036206	6	83357	170772407	c/g
rs2345689	6	83375	170772425	c/t
rs2345690	6	83381	170772431	c/t
rs2345691	6	83389	170772439	a/t
rs2345692	6	83443	170772493	a/g
rs3046306	6	83499	170772549	-/ggtg
rs4036207	6	83545	170772595	c/t
rs2345693	6	83566	170772616	c/t
rs2345694	6	83591	170772641	c/t
rs2345695	6	83619	170772669	g/t
rs2345696	6	83698	170772748	a/g
rs4036209	6	83780	170772830	g/t
rs2345697	6	83784	170772834	g/t
rs2881197	6	83826	170772876	g/t
rs2345698	6	83832	170772882	c/t
rs2345699	6	83852	170772902	c/t
rs2744640	6	86297	170775347	c/t
rs2744639	6	86315	170775365	g/t
rs2744638	6	86420	170775470	c/g
rs2744637	6	86460	170775510	c/g
rs2744636	6	86714	170775764	c/t
rs2744635	6	86718	170775768	c/t
rs2744634	6	86736	170775786	c/g
rs2744633	6	86753	170775803	c/t
rs2744632	6	86766	170775816	g/t
rs2744630	6	88162	170777212	c/g
rs2744629	6	88218	170777268	a/g
rs2744628	6	88246	170777296	a/g
rs2744627	6	88255	170777305	c/t
rs2977616	6	88309	170777359	g/t
rs2977617	6	88310	170777360	a/t
rs2744626	6	88471	170777521	a/g
rs2744625	6	88619	170777669	c/t
rs3115847	6	88904	170777954	c/t
rs2744623	6	89044	170778094	c/g
rs4036193	6	90531	170779581	-/aaaaa
rs4036194	6	90534	170779584	a/g
rs4036196	6	90613	170779663	c/g
Rs1042327	6	46252	170735302	c/t

Assay for Verifying and Allelotyping SNPs

[0230] The methods used to verify and allelotype the 109 proximal SNPs of Table 10 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 11 and Table 12, respectively.

TABLE 11

dbSNP rs#	Forward PCR primer	Reverse PCR primer
Rs1474555	ACGTTGGATGACATCAACTGAAGCCGACAG	ACGTTGGATGAATGGTGGAAATGTGATGAGA
Rs1474554	ACGTTGGATGATACACCTAGGACACCTCCA	ACGTTGGATGCAGAAGGAGATAAACCCAGC
rs10334	ACGTTGGATGAACAGTTTCCTCCCTGATGC	ACGTTGGATGCGGCTGGTGAAAGATGTCTT
rs10541	ACGTTGGATGACTATGCAGATCCGGAGTGC	ACGTTGGATGGTCCTTGACAGAGCCATG
Rs3823299	ACGTTGGATGCTCATGTGTACGAGGATTTG	ACGTTGGATGGTCTGGAAGGGTCTTTATTC
rs742348	ACGTTGGATGTGTGGATTTTCCAGTGCTCG	ACGTTGGATGCTGTACTTGAACCTCCCAAGC
Rs1474644	ACGTTGGATGGCAAGACAAGCATAATTGGG	ACGTTGGATGTAAAGGGCATTITGGCTTCC
Rs1474643	ACGTTGGATGTCTCCCAAATTAAGAGTGGC	ACGTTGGATGGATACCAAAGTCTACTTAC
Rs2056970	ACGTTGGATGTGGGACTACAGGAAGAGAAG	ACGTTGGATGCAAAACACAGACCTTCAGCC
Rs2223474	ACGTTGGATGCCAGGGTAAAGAAAAGATCC	ACGTTGGATGAGAGGCTTACCTCCTAAAAG
Rs2206284	ACGTTGGATGTCACATACTAGGTGGATCCC	ACGTTGGATGAAAGAGGAGAACACAGGATG
rs756519	ACGTTGGATGTCTAGAGACACCTGAGGTTG	ACGTTGGATGTGTTTCACTTCAGAGCCCTG
rs756518	ACGTTGGATGCCAGATTAGACTCTCTAAC	ACGTTGGATGAAATAGCTGAGCTGCCATTG
rs756517	ACGTTGGATGCTCGGTTGTTGACTCCTATC	ACGTTGGATGGCGGATGTTAAGAGTCAGAG
Rs1474642	ACGTTGGATGGGAGGTCATACATTAGCTTC	ACGTTGGATGTACCATCTGACACAATTCTC
Rs2038093	ACGTTGGATGGAGACAGAGTTTCACTCTTG	ACGTTGGATGTAATCACTTGAACCCAGGAG
Rs2038092	ACGTTGGATGTTACCTGAGGTCAGGAGTTT	ACGTTGGATGCCACACCCAGCTGATTTTTG
Rs2223473	ACGTTGGATGCCTTTATGTTATTGCTTTCC	ACGTTGGATGCAGGGAATTTAAGAATAGC
rs760909	ACGTTGGATGGGAAGAGGCAAGCTTAGTTC	ACGTTGGATGGCAGCATTAAACGAATGCCTG
Rs2076319	ACGTTGGATGGACATTTCAATGCCTTTG	ACGTTGGATGCCAACAGCAACTTAAAACTC
Rs3778589	ACGTTGGATGGCAAGAGAGAGAAAAGTTCC	ACGTTGGATGGTGTCTGTCTCCATTTTAC
Rs3800236	ACGTTGGATGAGAGAATGAGGCCCTATTTT	ACGTTGGATGCTCAGTCATTGTTCTTTTTT
Rs2206286	ACGTTGGATGTTGAGACGCTAACCCTCTAC	ACGTTGGATGAACATAGCCTCTGCTCTGTG
rs12717	ACGTTGGATGAAAATCGCAGCTGCAAAGGG	ACGTTGGATGAGACAGCAAGTGTCCGATCC
Rs2179373	ACGTTGGATGGAAGTGACCTATGCTCACAC	ACGTTGGATGAATGTCACTTCCGCCAGTTC
Rs3800235	ACGTTGGATGCTATGTGTTGATACCTCCAAG	ACGTTGGATGGCTTCATAAATGAAGTGAAC
Rs3823298	ACGTTGGATGGGTGGTTTCTTGCTTTGATG	ACGTTGGATGTTTTGTCCAGAGCATCTG
Rs2076318	ACGTTGGATGTCCGCCAAATTATTGTAGCC	ACGTTGGATGCTCAGTAGAAATGCATGGGC
Rs2235506	ACGTTGGATGTAACCATGTCAACTGTTCTC	ACGTTGGATGCCACCAACAATTTAGTAGG
Rs2072916	ACGTTGGATGACGCTGGAGTCACTAAGATG	ACGTTGGATGCAGATTAAGGCACAGGCATG
Rs3734763	ACGTTGGATGGCCTTTTGCTTTTCACTGTC	ACGTTGGATGTAAAGAGGCTGGACCTTCAG
Rs3177571	ACGTTGGATGGTCTGTTGTCAATATAGGTG	ACGTTGGATGACAAAAGTGTCCAGTGACAG
rs8770	ACGTTGGATGAATCCCTGTCACTGGACAC	ACGTTGGATGCCAAAAATAGAGGTGCAGAG
Rs3173219	ACGTTGGATGACATAACCACACTGGAGGTG	ACGTTGGATGCCTAGTTTTCAGACACGGTC
rs960744	ACGTTGGATGAAAGGCATGTCACAGTTCCC	ACGTTGGATGGCCCTCTGAGTCAGATAAAC
Rs2066954	ACGTTGGATGGAGGTTCTGGGTATAACTTTC	ACGTTGGATGCTACAAACCAAGTAAAGCTGATG
Rs2072917	ACGTTGGATGTGCTAGGCACTCACACTATC	ACGTTGGATGAGGCTTGGAAGTTCTCTCTG
Rs3173220	ACGTTGGATGTATCTGGGTTGACAAAGGCG	ACGTTGGATGACATAAGCAGGCTTGTGCAC
rs734249	ACGTTGGATGAGGTGGACACCAGCAGGGAA	ACGTTGGATGTCACCTCTGCACATGTCTTG

dbSNP rs#	Forward PCR primer	Reverse PCR primer
Rs2092310	ACGTTGGATGTTAGTCAGGTAAAGCGGGAC	ACGTTGGATGTCAGTGGAAGGCTGATCAAG
Rs2092309	ACGTTGGATGATCTAATTGCTTCCCCTCCC	ACGTTGGATGCAGCCTTCCACTGAATACAC
Rs1016536	ACGTTGGATGCCCCAAAAATTGGAGACAGG	ACGTTGGATGGGCTGTCATAATCGTGTGTC
Rs2235506	ACGTTGGATGAAGTGATTCTCCTGCCTCAG	ACGTTGGATGTGGTGAAACCCTGTCTCTAC
Rs2076998	ACGTTGGATGGCTCTGTGATTTTCGATGATG	ACGTTGGATGAGCTACTTCTTGCAGGAGTC
Rs2076997	ACGTTGGATGCAGAGCTTCCAAGTGTTTTC	ACGTTGGATGAAAGGAGTGCTTAAAGGAGC
Rs2345478	ACGTTGGATGCCTTCAACAAGTGCTGACAC	ACGTTGGATGATCCAGGCATTATTGCCAGC
Rs2021899	ACGTTGGATGGTTTTGTGGTGGATGATGGG	ACGTTGGATGAGAGTGCCCATATGGACAG
Rs2021898	ACGTTGGATGCGCAAGAACTCCTTGGATG	ACGTTGGATGCCAATTAAAGCCAAGGTCAC
Rs2345682	ACGTTGGATGATTGCGAAGAACTCCTTGG	ACGTTGGATGGGAAGAAATCTTACCAGAAC
Rs2345683	ACGTTGGATGATTGCGAAGAACTCCTTGG	ACGTTGGATGGGAAGAAATCTTACCAGAAC
Rs2881195	ACGTTGGATGATTGCGAAGAACTCCTTGG	ACGTTGGATGGGAAGAAATCTTACCAGAAC
Rs2345684	ACGTTGGATGATTGCGAAGAACTCCTTGG	ACGTTGGATGGGAAGAAATCTTACCAGAAC
Rs3046261	ACGTTGGATGCTCCACTCAGACATCAAAAG	ACGTTGGATGGTGACCTTGGCTTTAATTGG
Rs4083413	ACGTTGGATGGTGACCTTGGCTTTAATTGG	ACGTTGGATGCTCCACTCAGACATCAAAAG
Rs4083412	ACGTTGGATGGTGACCTTGGCTTTAATTGG	ACGTTGGATGCTCCACTCAGACATCAAAAG
Rs2345685	ACGTTGGATGGTCTGGTAAGATTCTTCC	ACGTTGGATGAGTCTTACAATAGATGACTG
Rs2021897	ACGTTGGATGGCAATTATTACAGAAGCCC	ACGTTGGATGTCCCACACAGTCATCTATTG
Rs4036211	ACGTTGGATGCCATTACAAGTTGGGCAGTT	ACGTTGGATGCTTTCTGATTCTTTTCTTCC
Rs4036212	ACGTTGGATGCTTTCTGATTCTTTTCTTCC	ACGTTGGATGCCATTACAAGTTGGGCAGTT
Rs4036213	ACGTTGGATGCCATTACAAGTTGGGCAGTT	ACGTTGGATGCTTTCTGATTCTTTTCTTCC
Rs2345686	ACGTTGGATGCCATTACAAGTTGGGCAGTT	ACGTTGGATGCTTTCTGATTCTTTTCTTCC
Rs4036214	ACGTTGGATGCCATTACAAGTTGGGCAGTT	ACGTTGGATGCTTTCTGATTCTTTTCTTCC
Rs4036215	ACGTTGGATGCTTTCTGATTCTTTTCTTCC	ACGTTGGATGCCATTACAAGTTGGGCAGTT
Rs2345687	ACGTTGGATGGGATTGTAAGGTGAGACTTG	ACGTTGGATGTTCTCCCATTAACAAGTTG
Rs2345688	ACGTTGGATGAGGGTCCCATCTAAGAATTC	ACGTTGGATGGGATTGTAAGGTGAGACTTG
Rs2881196	ACGTTGGATGAGGGTCCCATCTAAGAATTC	ACGTTGGATGGGATTGTAAGGTGAGACTTG
Rs3046288	ACGTTGGATGCCAATTGTAATGGGGAGGA	ACGTTGGATGCAGTTTTTACAGAGGGTCCC
Rs4036216	ACGTTGGATGCTTGTATGGGGAGGAAAAAA	ACGTTGGATGTTCTCATTTTAATCTGTCAG
Rs4036205	ACGTTGGATGCTTGTATGGGGAGGAAAAAA	ACGTTGGATGTTCTCATTTTAATCTGTCAG
Rs2092307	ACGTTGGATGCTTGTATGGGGAGGAAAAAA	ACGTTGGATGTTCTCATTTTAATCTGTCAG
Rs4036206	ACGTTGGATGGACCCTCTGTAAAACTGAC	ACGTTGGATGCCACTGCACCTCAAATCTTC
Rs2345689	ACGTTGGATGTTCCCTGAGTATCTCCCATG	ACGTTGGATGGGGACCCTCTGTAAAACTG
Rs2345690	ACGTTGGATGTTCCCTGAGTATCTCCCATG	ACGTTGGATGGGGACCCTCTGTAAAACTG
Rs2345691	ACGTTGGATGGCCACCTGTTGGAGATTAC	ACGTTGGATGGGGACCCTCTGTAAAACTG
Rs2345692	ACGTTGGATGTACATGGGAGATACTCAGGG	ACGTTGGATGCCACTGCACCTCAAATCTTC
Rs3046306	ACGTTGGATGGTATAACAAACCTTACCCTTG	ACGTTGGATGTAAAGAAAGAAGATTTGAGG
Rs4036207	ACGTTGGATGTATCAATGGAGAATGCGTGG	ACGTTGGATGGGGAGTTAACCAGCAAAAGC
Rs2345693	ACGTTGGATGTCGACAACAAGAAGAGAAGG	ACGTTGGATGCACATTAGACAAGGGTAAGG
Rs2345694	ACGTTGGATGCTACCTCTCTCGACAACAAG	ACGTTGGATGCTTAAAGTCCACGCATTCTCC
Rs2345695	ACGTTGGATGCGCATTCTCCATTGATAAGAC	ACGTTGGATGCCATTTAAAAGCTACCTCTC
Rs2345696	ACGTTGGATGCCTTACACAAGTGTAACCTC	ACGTTGGATGCCCCAAAATATAATGGTAGG
Rs4036209	ACGTTGGATGGGAACACAGTGATAAGACC	ACGTTGGATGGTTTTCACAACTTCGTTAGC
Rs2345697	ACGTTGGATGGTTTTCACAACTTCGTTAGC	ACGTTGGATGGCCACCCAAAATATAATGG
Rs2881197	ACGTTGGATGGCTGGAGGAAAAACAAGAAC	ACGTTGGATGCCTACCATTATATTTGGGG
Rs2345698	ACGTTGGATGCTGGAGGAAAAACAAGAACTC	ACGTTGGATGCATTATATTTGGGGTGGCAT
Rs2345699	ACGTTGGATGGCTGGAGGAAAAACAAGAAC	ACGTTGGATGGGGTGGCATATTTGGTCTT

dbSNP rs#	Forward PCR primer	Reverse PCR primer
Rs2744640	ACGTTGGATGGCAACAGCACTTAGTATGCC	ACGTTGGATGTGTGAAGCTGCAAATCTGGC
Rs2744639	ACGTTGGATGGCAACAGCACTTAGTATGCC	ACGTTGGATGTGTGAAGCTGCAAATCTGGC
Rs2744638	ACGTTGGATGAACCGTGGCAATACCACGTC	ACGTTGGATGTGGGTTTGGGCTGGATTGG
Rs2744637	ACGTTGGATGTGAGTTGACAGCCTCTGCTGG	ACGTTGGATGCACGTCAGTAAGGCAGAGAC
Rs2744636	ACGTTGGATGTCGGAGATGACATTGTCACC	ACGTTGGATGTTCCAGGGGTTACGTGTGTG
Rs2744635	ACGTTGGATGTGAGTCTGACTGTGTCACGG	ACGTTGGATGTCGGAGATGACATTGTCACC
Rs2744634	ACGTTGGATGCGTGTTCAGGGATTATATG	ACGTTGGATGGCACATAACGCTTGGAACTC
Rs2744633	ACGTTGGATGTATGAGTGTGACGGGTGTAG	ACGTTGGATGGCACATAACGCTTGGAACTC
Rs2744632	ACGTTGGATGTAGCTGCCTTCCACATCCAA	ACGTTGGATGTGTGACGGGTGTAGCGTTAG
Rs2744630	ACGTTGGATGGGGTTCAAATGCCTCTGATAG	ACGTTGGATGGGTCTAGGACAAGACCCATT
Rs2744629	ACGTTGGATGAACTTCCCTTAGCCAGTGG	ACGTTGGATGATCAGAGGCATTTGAACCCC
Rs2744628	ACGTTGGATGTTGACCTCAAATCATGTCAC	ACGTTGGATGTATCAGAGGCATTTGAACCCC
Rs2744627	ACGTTGGATGGGGTGGTTTATGTTCCACTG	ACGTTGGATGCCAGAACTAATGCTAGCTTC
Rs2977616	ACGTTGGATGTTCCACTGGCTAAGAGAAAG	ACGTTGGATGCCAGAACTAATGCTAGCTTC
Rs2977617	ACGTTGGATGCCAGAACTAATGCTAGCTTC	ACGTTGGATGTTCCACTGGCTAAGAGAAAG
Rs2744626	ACGTTGGATGACAGTGAAATTGTATTTCCG	ACGTTGGATGGCACAACCTTAAGAATCTCC
Rs2744625	ACGTTGGATGAGCAAAATCCACCTATGTCC	ACGTTGGATGCTGAATTTTGTCTCCAGTAC
Rs3115847	ACGTTGGATGTCGAGGCAGAGGCGTAGTA	ACGTTGGATGATAGGAATGACATGAACCCG
Rs2744623	ACGTTGGATGACGCGAGTCCGTAGGTGCTG	ACGTTGGATGAAGAGGCTGCTACCCAGAG
Rs4036193	ACGTTGGATGAGAGCAAGACTCCGTCTCAA	ACGTTGGATGACATGTCGCTTGATGTGTGC
Rs4036194	ACGTTGGATGACATGTCGCTTGATGTGTGC	ACGTTGGATGAGAGCAAGACTCCGTCTCAA
Rs4036196	ACGTTGGATGCCCCAGCGTTCATATTTGTC	ACGTTGGATGTCTGGCCAAATGGTCATACC
rs1042327	ACGTTGGATGAACTTCACATCACAGCTCCC	ACGTTGGATGCAGAAAGTTGGGTTTCCAGC

TABLE 12

dbSNP rs#	Extend Primer	Term Mix
rs1474555	TGAAGCCGACAGTGACACC	ACT
rs1474554	CCAATTTTGCACACCTCCAGCA	ACG
rs10334	CAGATCCGGAGTGCGTCC	CGT
rs10541	TCTCTCTCAGCCGCAGAA	CGT
rs3823299	GAGGATTTGTGATGAAAATACTA	ACG
rs742348	AATCCCCGTGTTGTTCAAGG	ACT
rs1474644	AAGGATGTTTCATCATAGTGTTTA	ACG
rs1474643	ACATGTTTATACATACTCATG	ACG
rs2056970	TTGGCAGCTTTTATAGGCCTC	ACT
rs2223474	AAGTCTCAAAAAGGTCCC	ACT
rs2206284	TAGGTGGATCCCTTTTCCC	ACG
rs756519	CAGAGCCCTGTTCTTTGATTT	ACG
rs756518	CAAAGGATGCTGTCTGGCC	ACG
rs756517	GTTCCATGAGCGTTTCTTTG	ACG
rs1474642	CTTCAGTTTCTTCATCACTTTC	ACT
rs2038093	TTTCACTCTTGTTGCCAGG	ACT
rs2038092	CCAACATGGTGAAACCCCATCT	ACT
rs2223473	TAGAATTAATAATAGACTTTGGGG	ACT

dbSNP rs#	Extend Primer	Term Mix
rs760909	GCAAGCTTAGTTCTAGGTCAG	CGT
rs2076319	TCACAATGCCTTTGTAATGATTT	ACT
rs3778589	GTTTTAGGAAGACTGCTCTGACAA	ACG
rs3800236	CTGAGAGCCAGCTGCAGTAA	CGT
rs2206286	CCTCGCCGGCTGGCATAA	ACT
rs12717	CCATCCCCAAGTCTCTGCCAG	ACT
rs2179373	TGACCTATGCTCACACTTCTCA	ACG
rs3800235	GTGTTGATACCTCCAAGTACATTT	CGT
rs3823298	CTTGATGAAATAGTCATCCAATA	ACT
rs2076318	TGAATTATCACCATCATCA	ACT
rs2235506	TGTTGCCAATAACAATCA	ACG
rs2072916	TGTGACAAGGGATTCCAC	ACG
rs3734763	CATCTGTAAGCAGGGCCGC	ACG
rs3177571	AAGACTGTGTAGCCTTCCTCTG	ACT
rs8770	GTAGACACTGTGTAAGCAATC	ACG
rs3173219	CACTGGAGGTGGAGAGCA	ACT
rs960744	CCCCATCAGACCTGGCTGT	ACT
rs2066954	TTACAATTTGAGCCTTGAGC	CGT
rs2072917	CTATCCCGACCCGAGAAAC	CGT
rs3173220	GCGATGAAACTGAACTGA	ACT
rs734249	CACCAGCAGGGAAGGTTTG	CGT
rs2092310	TTGAGGTGAGGGCTTCCAG	ACT
rs2092309	TCCCCTCCCCTATTGTTTAC	ACT
rs1016536	AAATTGGAGACAGGTCTCAGT	ACT
rs2235506	CTGGGAGTACAGGTGCGC	ACT
rs2076998	GTTTTTGTATAGTCTGCAGATGC	ACT
rs2076997	ATCCATTTTAATGGGTTGCTAGCT	ACT
rs2345478	ACAACGTACTTATTGGGCATA	ACT
rs2021899	CTTCTTGGAAGCTCTTCCCA	ACT
rs2021898	TTGGATGGGGTTAATGGCAG	ACG
rs2345682	GTTAATGGCAGCTGTATTTTCTG	ACT
rs2345683	GGCAGCTGTATTTTCTGTGA	ACT
rs2881195	CAGCTGTATTTTCTGTGACCT	ACG
rs2345684	GCAGCTGTATTTTCTGTGACCTT	ACT
rs3046261	GAAAACATTTGAGATACTGAAGAT	ACT
rs4083413	TTCTTTATCTTCAGTATCTCAA	ACT
rs4083412	TCTTCAGTATCTCAAATGTTTTCA	ACG
rs2345685	CAACTTTTGATGTCTGAGTGGA	ACT
rs2021897	ATTATTTACAGAAGCCCTATTCA	ACT
rs4036211	TTCCAAACAAAAGCTACCATGCA	ACT
rs4036212	AAATAATTGCATGGTAGCTTTTG	CGT
rs4036213	ACAACACTTTTGATGTTATTTCC	CGT
rs2345686	ACAATCCAAAAATCACATTCCTA	ACT
rs4036214	GTCTCACCTTACAATCCAAAAAT	CGT

dbSNP rs#	Extend Primer	Term Mix
rs4036215	AATGTGATTTTTGGATTGTAAGG	ACT
rs2345687	AAGGTGAGACTTGTTTAGCTTT	ACT
rs2345688	TCCTCCCCATTACAAGTTGGGCA	ACT
rs2881196	TTTCCTCCCCATTACAAGTTGG	ACT
rs3046288	TAATGGGGAGGAAAAAAATTTTCT	ACT
rs4036216	ATGTTTTTGAATTCTTAGATGG	ACT
rs4036205	GTTTTTGAATTCTTAGATGGGAC	ACT
rs2092307	TGGAATTCTTAGATGGGACCC	ACT
rs4036206	ACTGACAGATTAAAATGAGAAAAA	ACT
rs2345689	TCCCATGTATCCATAAGGTATAC	ACT
rs2345690	GTATCTCCCATGTATCCATAAG	ACT
rs2345691	CCCTGAGTATCTCCCATGTA	CGT
rs2345692	TCTCCAACAGGTGGCTTTCA	ACT
rs3046306	TTGCTGGTTAACTCCCCACT	CGT
rs4036207	GCGTGGACTTAAGTCTGTATAAC	ACT
rs2345693	AGAGTCTTATCAATGGAGAATGC	ACT
rs2345694	GAAGAGAAGGATAACTAAATCACT	ACT
rs2345695	ATTTAGTTATCCTTCTCTTCTTG	ACT
rs2345696	ACACAAGTGTAACCTTCTACTCT	ACT
rs4036209	GGAAACCAGAATATGCCACC	CGT
rs2345697	AGCCAAAGGGACATATTTTGTGGT	ACT
rs2881197	GGAACACAGTGTATAAGACCAAA	CGT
rs2345698	CGGTGGAACACAGTGTATAAG	ACT
rs2345699	AAAACAAGAACTCTTTTCATTGCC	ACT
rs2744640	TTTATCTCCAGTTCCCCAGC	ACG
rs2744639	AGCACTTAGTATGCCTTCTCCTT	ACT
rs2744638	TGGCAATACCACGTCAGTAAG	ACT
rs2744637	GCTGGGCTGGGTTTGGGCTG	ACT
rs2744636	ACCCGTCACACTCATATAATCCC	ACG
rs2744635	ACACATGCGTGTTCCAGGG	ACT
rs2744634	GGGATTATATGAGTGTGACGG	ACT
rs2744633	GGGTGTAGCGTTAGGTGAC	ACT
rs2744632	GCGCACATAACGCTTGGAAC	ACT
rs2744630	CGTGTTAAACTCATGGCCAAAC	ACT
rs2744629	ATAAACCACCCTGGAGTTCAT	ACT
rs2744628	TTGAAGAAAACTTTCCCTTAGCCA	ACT
rs2744627	GTTTATGTTCCACTGGCTAAG	ACT
rs2977616	TTGAGGTCAAACATTAATATCAAG	ACT
rs2977617	CTAGCTTCTCAATCTTTTGAGTT	CGT
rs2744626	GTGAAATTGTATTTCCGGATTTC	ACT
rs2744625	TCCTGAACACTTATCCACTTTAC	ACT
rs3115847	CCAGGGCTGGAGGGGCC	ACT
rs2744623	GGTGCTGGCGGGAGCGAGAGT	ACT
rs4036193	GACTCCGTCTCAAAAAAAAAAAAAA	ACT

dbSNP rs#	Extend Primer	Term Mix
rs4036194	CTTGATGTGTGCTTCAGGGTA	ACG
rs4036196	CAGTGCAAGTAAAGAGCCTTA	ACT
rs1042327	CATCACAGCTCCCCACCAT	ACT

Genetic Analysis

[0231] Allelotyping results from the discovery cohort are shown for cases and controls in Table 13. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where “AF” is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs1474555 has the following case and control allele frequencies: case A1 (C) = 0.64; case A2 (T) = 0.36; control A1 (C) = 0.70; and control A2 (T) = 0.30, where the nucleotide is provided in paranthesis. Some SNPs are labeled “untyped” because of failed assays.

TABLE 13

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
Rs1474555	229	170689279	C/T	0.36	0.30	0.024
Rs1474554	6310	170695360	A/G	0.48	0.43	0.058
rs10334	11840	170700890	G/T			
rs10541	11870	170700920	A/T			
Rs3823299	12064	170701114	A/G	0.45	0.41	0.125
Rs742348	13392	170702442	C/G	0.46	0.44	0.275
Rs1474644	16354	170705404	A/G	0.75	0.77	0.270
Rs1474643	16559	170705609	C/T	0.45	0.40	0.042
Rs2056970	16935	170705985	A/G	0.36	0.33	0.242
Rs2223474	17616	170706666	C/T	0.42	0.46	0.140
Rs2206284	17737	170706787	C/T	0.37	0.35	0.493
rs756519	18321	170707371	C/T			
Rs756518	18453	170707503	A/G	0.49	0.53	0.133
Rs756517	18811	170707861	C/T			
Rs1474642	20020	170709070	C/T	0.12	0.12	0.904
Rs2038093	21662	170710712	C/G			
Rs2038092	23197	170712247	C/G			
Rs2223473	23446	170712496	G/T	0.42	0.45	0.296
Rs760909	24339	170713389	G/T	0.49	0.52	0.255
Rs2076319	25504	170714554	A/G	0.43	0.46	0.219
Rs3778589	27174	170716224	A/G	0.49	0.54	0.081
Rs3800236	28008	170717058	A/T	0.47	0.50	0.319
Rs2206286	29294	170718344	C/T	0.81	0.82	0.831
rs12717	29759	170718809	C/G	0.52	0.57	0.081
rs2179373	30832	170719882	A/G	0.58	0.62	0.089
rs3800235	44512	170733562	A/C	0.60	0.64	0.077
rs3823298	44850	170733900	C/G	0.44	0.38	0.022
rs2076318	45884	170734934	A/G	0.41	0.45	0.109
rs2235506	46345	170735395	C/T	0.68	0.66	0.320
rs2072916	48589	170737639	A/G	0.48	0.51	0.192
rs3734763	53371	170742421	A/G	0.50	0.54	0.142

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs3177571	53911	170742961	G/T			
rs8770	53990	170743040	A/G			
rs3173219	55152	170744202	C/G	0.49	0.53	0.056
rs960744	55667	170744717	C/T	0.39	0.35	0.179
rs2066954	58952	170748002	A/C	0.37	0.32	0.057
rs2072917	59315	170748365	G/T	0.46	0.42	0.153
rs3173220	60029	170749079	A/G			
rs734249	61477	170750527	A/C	0.48	0.40	0.022
rs2092310	62988	170752038	C/T			
rs2092309	63090	170752140	C/G	0.43	0.47	0.165
rs1016536	64021	170753071	A/C	0.10	0.10	0.985
rs2235506	65685	170754735	C/T			
rs2076998	70220	170759270	A/G			
rs2076997	70323	170759373	A/C	0.90	0.90	0.814
rs2345478	70959	170760009	A/C	0.09	0.09	0.947
rs2021899	73436	170762486	C/G	0.46	0.43	0.218
rs2021898	82945	170771995	A/G			
rs2345682	82958	170772008	G/T			
rs2345683	82961	170772011	C/G	0.28	0.34	0.019
rs2881195	82964	170772014	C/T			
rs2345684	82965	170772015	G/T			
rs3046261	83006	170772056	-JCTTT			
rs4083413	83025	170772075	C/T			
rs4083412	83034	170772084	A/G			
rs2345685	83074	170772124	G/T	0.71	0.71	0.835
rs2021897	83132	170772182	G/T			
rs4036211	83155	170772205	C/T			
rs4036212	83172	170772222	A/T			
rs4036213	83174	170772224	G/T			
rs2345686	83206	170772256	C/T			
rs4036214	83216	170772266	G/T			
rs4036215	83234	170772284	G/T			
rs2345687	83252	170772302	A/G	0.55	0.50	0.085
rs2345688	83260	170772310	A/C	0.53	0.52	0.958
rs2881196	83263	170772313	A/C			
rs3046288	83296	170772346	-JAT			
rs4036216	83319	170772369	A/G			
rs4036205	83322	170772372	C/G			
rs2092307	83324	170772374	A/C			
rs4036206	83357	170772407	C/G			
rs2345689	83375	170772425	C/T			
rs2345690	83381	170772431	C/T			
rs2345691	83389	170772439	A/T			
rs2345692	83443	170772493	A/G			
rs3046306	83499	170772549	-GGTG	0.42	0.43	0.761
rs4036207	83545	170772595	C/T			
rs2345693	83566	170772616	C/T			
rs2345694	83591	170772641	C/T			
rs2345695	83619	170772669	G/T			
rs2345696	83698	170772748	A/G			
rs4036209	83780	170772830	G/T	0.79	0.73	0.156
rs2345697	83784	170772834	G/T			
rs2881197	83826	170772876	G/T			
rs2345698	83832	170772882	C/T			
rs2345699	83852	170772902	C/T			
rs2744640	86297	170775347	C/T	0.53	0.53	0.973
rs2744639	86315	170775365	G/T	0.40	0.40	0.789

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2744638	86420	170775470	C/G	0.39	0.39	0.941
rs2744637	86460	170775510	C/G	0.40	0.42	0.497
rs2744636	86714	170775764	C/T	0.76	0.73	0.271
rs2744635	86718	170775768	C/T	0.03	0.02	0.425
rs2744634	86736	170775786	C/G	0.96	0.94	0.436
rs2744633	86753	170775803	C/T	0.14	0.16	0.409
rs2744632	86766	170775816	G/T	0.80	0.83	0.217
rs2744630	88162	170777212	C/G			
rs2744629	88218	170777268	A/G	0.80	0.80	0.978
rs2744628	88246	170777296	A/G	0.71	0.67	0.206
rs2744627	88255	170777305	C/T	0.32	0.30	0.335
rs2977616	88309	170777359	G/T			
rs2977617	88310	170777360	A/T			
rs2744626	88471	170777521	A/G			
rs2744625	88619	170777669	C/T			
rs3115847	88904	170777954	C/T			
rs2744623	89044	170778094	C/G			
rs4036193	90531	170779581	-/AAAAA			
rs4036194	90534	170779584	A/G			
rs4036196	90613	170779663	C/G			
rs1042327	46252	170735302	C/T	0.45	0.39	0.028

[0232] The chromosome 6 proximal SNPs were also allelotyped in the replication cohorts using the methods described herein and the primers provided in Tables 11 and 12. The replication allelotyping results for replication cohort #1 and replication cohort #2 are provided in Tables 14 and 15, respectively.

TABLE 14

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1474555	229	170689279	C/T	0.37	0.27	0.004
rs1474554	6310	170695360	A/G	0.50	0.42	0.020
rs10334	11840	170700890	G/T			
rs10541	11870	170700920	A/T			
rs3823299	12064	170701114	A/G	0.45	0.40	0.080
rs742348	13392	170702442	C/G	0.47	0.41	0.075
rs1474644	16354	170705404	A/G	0.75	0.79	0.231
rs1474643	16559	170705609	C/T	0.46	0.39	0.028
rs2056970	16935	170705985	A/G	0.38	0.33	0.129
rs2223474	17616	170706666	C/T	0.41	0.48	0.052
rs2206284	17737	170706787	C/T	0.37	0.34	0.342
rs756519	18321	170707371	C/T			
rs756518	18453	170707503	A/G	0.48	0.56	0.013
rs756517	18811	170707861	C/T			
rs1474642	20020	170709070	C/T	0.10	0.13	0.277
rs2038093	21662	170710712	C/G			
rs2038092	23197	170712247	C/G			
rs2223473	23446	170712496	G/T	0.42	0.48	0.070
rs760909	24339	170713389	G/T	0.47	0.54	0.077
rs2076319	25504	170714554	A/G	0.41	0.49	0.017
rs3778589	27174	170716224	A/G	0.50	0.57	0.035
rs3800236	28008	170717058	A/T	0.47	0.52	0.126
rs2206286	29294	170718344	C/T	0.80	0.80	0.952

dbSNP rs#	Position in SEQ ID NO: I	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs12717	29759	170718809	C/G	0.53	0.59	0.059
rs2179373	30832	170719882	A/G	0.57	0.64	0.025
rs3800235	44512	170733562	A/C	0.59	0.65	0.065
rs3823298	44850	170733900	C/G	0.46	0.36	0.003
rs2076318	45884	170734934	A/G	0.40	0.47	0.017
rs2235506	46345	170735395	C/T	0.68	0.65	0.434
rs2072916	48589	170737639	A/G	0.47	0.54	0.026
rs3734763	53371	170742421	A/G	0.49	0.56	0.052
rs3177571	53911	170742961	G/T			
rs8770	53990	170743040	A/G			
rs3173219	55152	170744202	C/G	0.49	0.55	0.069
rs960744	55667	170744717	C/T	0.39	0.34	0.131
rs2066954	58952	170748002	A/C	0.36	0.31	0.096
rs2072917	59315	170748365	G/T	0.46	0.41	0.070
rs3173220	60029	170749079	A/G			
rs734249	61477	170750527	A/C	0.37	NA	0.484
rs2092310	62988	170752038	C/T			
rs2092309	63090	170752140	C/G	0.43	0.49	0.102
rs1016536	64021	170753071	A/C	0.08	0.11	0.277
rs2235506	65685	170754735	C/T			
rs2076998	70220	170759270	A/G			
rs2076997	70323	170759373	A/C	0.89	0.91	0.655
rs2345478	70959	170760009	A/C	0.08	0.09	0.660
rs2021899	73436	170762486	C/G	0.48	0.42	0.081
rs2021898	82945	170771995	A/G			
rs2345682	82958	170772008	G/T			
rs2345683	82961	170772011	C/G	0.32	0.39	0.046
rs2881195	82964	170772014	C/T			
rs2345684	82965	170772015	G/T			
rs3046261	83006	170772056	-/CTTT			
rs4083413	83025	170772075	C/T			
rs4083412	83034	170772084	A/G			
rs2345685	83074	170772124	G/T	0.69	0.70	0.772
rs2021897	83132	170772182	G/T			
rs4036211	83155	170772205	C/T			
rs4036212	83172	170772222	A/T			
rs4036213	83174	170772224	G/T			
rs2345686	83206	170772256	C/T			
rs4036214	83216	170772266	G/T			
rs4036215	83234	170772284	G/T			
rs2345687	83252	170772302	A/G	0.62	NA	NA
rs2345688	83260	170772310	A/C	0.46	0.49	0.383
rs2881196	83263	170772313	A/C			
rs3046288	83296	170772346	-/AT			
rs4036216	83319	170772369	A/G			
rs4036205	83322	170772372	C/G			
rs2092307	83324	170772374	A/C			
rs4036206	83357	170772407	C/G			
rs2345689	83375	170772425	C/T			
rs2345690	83381	170772431	C/T			
rs2345691	83389	170772439	A/T			
rs2345692	83443	170772493	A/G			
rs3046306	83499	170772549	-/GGTG	0.39	0.40	0.729
rs4036207	83545	170772595	C/T			
rs2345693	83566	170772616	C/T			
rs2345694	83591	170772641	C/T			
rs2345695	83619	170772669	G/T			

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2345696	83698	170772748	A/G			
rs4036209	83780	170772830	G/T	0.79	0.73	0.156
rs2345697	83784	170772834	G/T			
rs2881197	83826	170772876	G/T			
rs2345698	83832	170772882	C/T			
rs2345699	83852	170772902	C/T			
rs2744640	86297	170775347	C/T	0.49	0.51	0.583
rs2744639	86315	170775365	G/T	0.45	0.43	0.745
rs2744638	86420	170775470	C/G	0.38	0.38	0.852
rs2744637	86460	170775510	C/G	0.35	0.40	0.216
rs2744636	86714	170775764	C/T	0.71	0.73	0.482
rs2744635	86718	170775768	C/T	0.05	0.03	0.195
rs2744634	86736	170775786	C/G	0.93	0.92	0.601
rs2744633	86753	170775803	C/T	0.19	0.20	0.681
rs2744632	86766	170775816	G/T	0.85	0.90	0.070
rs2744630	88162	170777212	C/G			
rs2744629	88218	170777268	A/G	0.78	0.79	0.891
rs2744628	88246	170777296	A/G	0.68	0.67	0.766
rs2744627	88255	170777305	C/T	0.32	0.30	0.636
rs2977616	88309	170777359	G/T			
rs2977617	88310	170777360	A/T			
rs2744626	88471	170777521	A/G			
rs2744625	88619	170777669	C/T			
rs3115847	88904	170777954	C/T			
rs2744623	89044	170778094	C/G			
rs4036193	90531	170779581	-IAAAAA			
rs4036194	90534	170779584	A/G			
rs4036196	90613	170779663	C/G			
rs1042327	46252	170735302	C/T	0.46	0.37	0.004

TABLE 15

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1474555	229	170689279	C/T	0.35	0.36	0.770
rs1474554	6310	170695360	A/G	0.45	0.44	0.873
rs10334	11840	170700890	G/T			
rs10541	11870	170700920	A/T			
rs3823299	12064	170701114	A/G	untyped	0.43	NA
rs742348	13392	170702442	C/G	0.45	0.47	0.600
rs1474644	16354	170705404	A/G	0.74	0.75	0.775
rs1474643	16559	170705609	C/T	0.43	0.41	0.614
rs2056970	16935	170705985	A/G	0.33	0.33	0.978
rs2223474	17616	170706666	C/T	0.44	0.43	0.944
rs2206284	17737	170706787	C/T	0.36	0.37	0.901
rs756519	18321	170707371	C/T			
rs756518	18453	170707503	A/G	0.50	0.47	0.453
rs756517	18811	170707861	C/T			
rs1474642	20020	170709070	C/T	0.15	0.11	0.147
rs2038093	21662	170710712	C/G			
rs2038092	23197	170712247	C/G			
rs2223473	23446	170712496	G/T	0.43	0.40	0.408
rs760909	24339	170713389	G/T	0.51	0.48	0.506
rs2076319	25504	170714554	A/G	0.44	0.40	0.264
rs3778589	27174	170716224	A/G	0.49	0.48	0.910

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs3800236	28008	170717058	A/T	0.48	0.46	0.670
rs2206286	29294	170718344	C/T	0.83	0.84	0.685
rs12717	29759	170718809	C/G	0.51	0.53	0.726
rs2179373	30832	170719882	A/G	0.59	0.58	0.880
rs3800235	44512	170733562	A/C	0.60	0.62	0.632
rs3823298	44850	170733900	C/G	0.41	0.41	0.945
rs2076318	45884	170734934	A/G	0.43	0.42	0.636
rs2235506	46345	170735395	C/T	0.69	0.67	0.594
rs2072916	48589	170737639	A/G	0.49	0.46	0.399
rs3734763	53371	170742421	A/G	0.51	0.51	0.888
rs3177571	53911	170742961	G/T			
rs8770	53990	170743040	A/G			
rs3173219	55152	170744202	C/G	0.48	0.51	0.493
rs960744	55667	170744717	C/T	0.38	0.37	0.738
rs2066954	58952	170748002	A/C	0.37	0.34	0.378
rs2072917	59315	170748365	G/T	0.45	0.45	0.982
rs3173220	60029	170749079	A/G			
rs734249	61477	170750527	A/C	0.46	0.02	
rs2092310	62988	170752038	C/T			
rs2092309	63090	170752140	C/G	0.43	0.44	0.891
rs1016536	64021	170753071	A/C	0.13	0.09	0.173
rs2235506	65685	170754735	C/T			
rs2076998	70220	170759270	A/G			
rs2076997	70323	170759373	A/C	0.92	0.89	0.256
rs2345478	70959	170760009	A/C	0.11	0.10	0.545
rs2021899	73436	170762486	C/G	0.44	0.45	0.797
rs2021898	82945	170771995	A/G			
rs2345682	82958	170772008	G/T			
rs2345683	82961	170772011	C/G	0.23	0.26	0.407
rs2881195	82964	170772014	C/T			
rs2345684	82965	170772015	G/T			
rs3046261	83006	170772056	-/CTTT			
rs4083413	83025	170772075	C/T			
rs4083412	83034	170772084	A/G			
rs2345685	83074	170772124	G/T	0.74	0.71	0.533
rs2021897	83132	170772182	G/T			
rs4036211	83155	170772205	C/T			
rs4036212	83172	170772222	A/T			
rs4036213	83174	170772224	G/T			
rs2345686	83206	170772256	C/T			
rs4036214	83216	170772266	G/T			
rs4036215	83234	170772284	G/T			
rs2345687	83252	170772302	A/G	0.47	0.50	0.457
rs2345688	83260	170772310	A/C	0.61	0.58	0.434
rs2881196	83263	170772313	A/C			
rs3046288	83296	170772346	-/AT			
rs4036216	83319	170772369	A/G			
rs4036205	83322	170772372	C/G			
rs2092307	83324	170772374	A/C			
rs4036206	83357	170772407	C/G			
rs2345689	83375	170772425	C/T			
rs2345690	83381	170772431	C/T			
rs2345691	83389	170772439	A/T			
rs2345692	83443	170772493	A/G			
rs3046306	83499	170772549	-/GGTG			
rs4036207	83545	170772595	C/T			
rs2345693	83566	170772616	C/T			

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2345694	83591	170772641	C/T			
rs2345695	83619	170772669	G/T			
rs2345696	83698	170772748	A/G			
rs4036209	83780	170772830	G/T			
rs2345697	83784	170772834	G/T			
rs2881197	83826	170772876	G/T			
rs2345698	83832	170772882	C/T			
rs2345699	83852	170772902	C/T			
rs2744640	86297	170775347	C/T	0.57	0.55	0.595
rs2744639	86315	170775365	G/T	0.35	0.34	0.752
rs2744638	86420	170775470	C/G	0.41	0.40	0.793
rs2744637	86460	170775510	C/G	0.47	0.46	0.836
rs2744636	86714	170775764	C/T	0.83	NA	
rs2744635	86718	170775768	C/T			
rs2744634	86736	170775786	C/G	untyped	0.97	NA
rs2744633	86753	170775803	C/T	0.09	0.10	0.691
rs2744632	86766	170775816	G/T	0.74	0.72	0.529
rs2744630	88162	170777212	C/G			
rs2744629	88218	170777268	A/G	0.81	0.81	0.959
rs2744628	88246	170777296	A/G	0.74	NA	
rs2744627	88255	170777305	C/T	0.33	0.29	0.341
rs2977616	88309	170777359	G/T			
rs2977617	88310	170777360	A/T			
rs2744626	88471	170777521	A/G			
rs2744625	88619	170777669	C/T			
rs3115847	88904	170777954	C/T			
rs2744623	89044	170778094	C/G			
rs4036193	90531	170779581	-/AAAAA			
rs4036194	90534	170779584	A/G			
rs4036196	90613	170779663	C/G			
rs1042327	46252	170735302	C/T	0.42	0.43	0.880

[0233] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1A for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1A can be determined by consulting Table 13. For example, the left-most X on the left graph is at position 170689279. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0234] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial

regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than 10^{-8} were truncated at that value.

[0235] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

Example 5

ELP3 Region Proximal SNPs

[0236] It has been discovered that SNP rs1563055 in elongation protein 3 homolog (*ELP3*) is associated with occurrence of osteoarthritis in subjects.

[0237] Thirty-three additional allelic variants proximal to rs1563055 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 16. The chromosome positions provided in column four of Table 16 are based on Genome "Build 34" of NCBI's GenBank.

TABLE 16

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs1000658	8	211	27927511	c/t
rs1984880	8	473	27927773	c/t
rs999112	8	1536	27928836	c/t
rs735880	8	5639	27932939	c/t
rs2045029	8	17186	27944486	a/g
rs2045028	8	17335	27944635	c/t
rs1947384	8	25029	27952329	c/g
rs1947385	8	25111	27952411	c/t
rs1901744	8	28811	27956111	a/g
rs1901745	8	28863	27956163	a/t
rs971882	8	30809	27958109	a/c
rs1377338	8	40985	27968285	a/c
rs2305452	8	45147	27972447	c/t
rs2305451	8	45282	27972582	a/g
rs2123472	8	46168	27973468	g/t
rs2167768	8	46328	27973628	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs1563055	8	49077	27976377	a/g
rs2290371	8	51925	27979225	c/t
rs2290370	8	52141	27979441	a/g
rs2290369	8	52168	27979468	c/t
rs2874904	8	60852	27988152	c/t
rs3213997	8	62468	27989768	a/g
rs3213998	8	65572	27992872	g/t
rs1530929	8	79089	28006389	a/c
rs1000275	8	79541	28006841	c/t
rs1000274	8	79790	28007090	c/t
rs3757896	8	90843	28018143	a/g
rs3757895	8	90978	28018278	c/t
rs3757894	8	91052	28018352	c/g
rs3757893	8	91131	28018431	a/g
rs3757892	8	91132	28018432	c/t
rs3757891	8	94439	28021739	a/g
rs3757890	8	94621	28021921	a/t

Assay for Verifying and Allelotyping SNPs

[0238] The methods used to verify and allelotype the 33 proximal SNPs of Table 16 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 17 and Table 18, respectively.

TABLE 17

dbSNP rs#	Forward PCR primer	Reverse PCR primer
Rs1000658	ACGTTGGATGTTCTCAAAAAAGAAACACAT	ACGTTGGATGGGGTTATCAGTTTGAGATTC
Rs1984880	ACGTTGGATGCCATTTGCCAATTCCTGTGG	ACGTTGGATGATGGGCTGAAATGTATCCCC
rs999112	ACGTTGGATGCTAAGCACATGCCTTTCTTG	ACGTTGGATGCTATTTTCTACTGGGAGATG
rs735880	ACGTTGGATGTGCCTTCATTCTCCAACCAC	ACGTTGGATGAACAGAGTGAGACCCATCTG
Rs2045029	ACGTTGGATGAGTCATTGCTAGCTTTCTGG	ACGTTGGATGGGGACTTTAGGGAAGTTATAG
Rs2045028	ACGTTGGATGAGCTTGAGTGAGCCGAGAT	ACGTTGGATGTGAGACAGAGTCTTGCTCTG
Rs1947384	ACGTTGGATGATTCTCCACCGAGAAACCAG	ACGTTGGATGTTGTGGCAGCAAGAAGGAAC
Rs1947385	ACGTTGGATGAAATTTCAACAGTCAACAAT	ACGTTGGATGGTCAGTTTTGAAAAGTATC
Rs1901744	ACGTTGGATGCCTTGATTGAAGAGTAAAGC	ACGTTGGATGATCAAATATTCCTCATCCCC
Rs1901745	ACGTTGGATGCTTCTGCCTTTACCTGTGTC	ACGTTGGATGAAATGAAGCAGCACTCACAG
rs971882	ACGTTGGATGAAGCCCTAATCATTGGTACG	ACGTTGGATGGATGGGTGCTAAAAAGACAC
Rs1377338	ACGTTGGATGCCACATATCTACACATCAAG	ACGTTGGATGAGGGAGATAGGTGGTTAAAG
Rs2305452	ACGTTGGATGCCGTGTTGCAACTAACAGGG	ACGTTGGATGAGACGTTCCCATCCTCCATC
Rs2305451	ACGTTGGATGGCAGAGCCACCAGAGATAAA	ACGTTGGATGTTTTACGACAGGCGGGATTG
Rs2123472	ACGTTGGATGCACTTAGAATTGTTGCTTGG	ACGTTGGATGGCTGTATCTGTGACCTCAA
Rs2167768	ACGTTGGATGGAATCAACATGACTTGGTGAC	ACGTTGGATGATCTCACTCTAACTTGCTCC
rs1563055	ACGTTGGATGAGTTCTTTCTCCTCACATTG	ACGTTGGATGCCCTTTAGAAGCACATACTC
Rs2290371	ACGTTGGATGATCCTCTTGGTAGCTTGTC	ACGTTGGATGCTGTCTTGGTTTTACCCCTG

dbSNP rs#	Forward PCR primer	Reverse PCR primer
Rs2290370	ACGTTGGATGCAACCTCTACCTCACTACAC	ACGTTGGATGATGAGGTATCGACACACTGG
Rs2290369	ACGTTGGATGACACACTGGGTATCTGTTCT	ACGTTGGATGTCAGAATCCCCAACCTCTAC
Rs2874904	ACGTTGGATGAAATTCCAGGCTGGGTACAG	ACGTTGGATGTGCTGACCTTAAGTGATCCG
Rs3213997	ACGTTGGATGGGTTGGCTAGAAGAGAAAAA	ACGTTGGATGTACAGTCCTTTTGAAACTAC
Rs3213998	ACGTTGGATGACAGTTTGTGACATAGTAG	ACGTTGGATGAGGCTGAAAAGACATTTCATG
Rs1530929	ACGTTGGATGGGCTTTCACTATATTTCTC	ACGTTGGATGGAATACAGTAAGCCTATGGG
Rs1000275	ACGTTGGATGAACCCAGAAAGCAAAAAGC	ACGTTGGATGCACGCTTGCTAACTTAATGG
Rs1000274	ACGTTGGATGGCCTAAGACAGGATCCAAAC	ACGTTGGATGTTACTGCGTGCCTTAGTACC
Rs3757896	ACGTTGGATGCCTTCAAGCAAGTCAGTTAC	ACGTTGGATGCAGAACTGTGTGACTGATC
Rs3757895	ACGTTGGATGAAAATCATTGGCCAAACTGC	ACGTTGGATGCTCCTTAGTATTCTTAGGTG
Rs3757894	ACGTTGGATGAGAAGGGTTGAACAACAAGG	ACGTTGGATGCACCTAAGAATACTAAGGAG
Rs3757893	ACGTTGGATGCCCTTGTGTTCAACCCTTC	ACGTTGGATGCTGCATGTGGATACCTACAC
Rs3757892	ACGTTGGATGTCCTGCATGTGGATACCTAC	ACGTTGGATGCCCTTGTGTTCAACCCTTC
Rs3757891	ACGTTGGATGATGGGCCAATTCTCCATAGG	ACGTTGGATGAGGCCTGTTAAGGAAACCTG
Rs3757890	ACGTTGGATGCAGGTGGATGTAGGCTTAAG	ACGTTGGATGGCACCCTGCCTCTTGTTTT

TABLE 18

dbSNP rs#	Extend Primer	Term Mix
rs1000658	AATTGACAATGTTGGGACTGTT	ACG
rs1984880	TGTGGTGTAATAGGAGTTAGTGG	ACT
rs999112	GCACATGCCTTTCTTGGAAGT	ACG
rs735880	AACCTTTACTTGTACTACATGC	ACG
rs2045029	GCTAGCTTTCTGGTAATGAAAAT	ACT
rs2045028	GATCGCACCCTGCACTCCAG	ACG
rs1947384	ATAGCGGCAGTCCAAAAAGC	ACT
rs1947385	TTCAACAGTCAACAATGAAACC	ACT
rs1901744	ATAGTCAAGTATGCAAATGAAGC	ACT
rs1901745	CCTTTACCTGTGTCTTCCCT	CGT
rs971882	CCTAATCATTGGTACGGTCTCA	ACT
rs1377338	AGTATTAGCTCAAATATCACATTG	ACT
rs2305452	CAGGGTAGCAGGCGGCC	ACG
rs2305451	CCACAACTCAGACCACGG	ACT
rs2123472	CAGTTAATGTCAAGAAGCATAG	ACT
rs2167768	ACATGACTTGGTGACAGAAGAA	ACT
rs1563055	TTCTCCTCACATTGTTTCTACT	ACG
rs2290371	GGTAGCTTGTCTTAAATAACCGT	ACT
rs2290370	GGAGCAGGGACTTCTGCCA	ACT
rs2290369	AGTCCCTGCTCCATGTGAC	ACT
rs2874904	GGCTAACGCCTGTAATCCCA	ACT
rs3213997	AGAAAAATATTGTTATGCCACA	ACG
rs3213998	TAGTATTCTCAAATAGAGAGATTC	ACT
rs1530929	TTTCCTCTTTCCAGAATTGTATT	ACT
rs1000275	ATGAGAATATCCTAGAATGAGGCA	ACG

dbSNP rs#	Extend Primer	Term Mix
rs1000274	GAATCATCAGGTCCTGTGCC	ACG
rs3757896	TAATTCTCCTTAAGTAGTTAATTC	ACT
rs3757895	TTGGCCAAACTGCAGGATCT	ACT
rs3757894	AAGGGCCACACAAGCAATTTCAA	ACT
rs3757893	CCAAAGGACATTAGGTGGTG	ACG
rs3757892	TGTGGATACCTACACTGCTC	ACG
rs3757891	AGGATAAGTGTAAACGGGGTC	ACT
rs3757890	AGTGACACTCTTACTTCACAC	CGT

Genetic Analysis

[0239] Allelotyping results from the discovery cohort are shown for cases and controls in Table 19. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where “AF” is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs1000658 has the following case and control allele frequencies: case A1 (C) = 0.36; case A2 (T) = 0.64; control A1 (C) = 0.37; and control A2 (T) = 0.63, where the nucleotide is provided in paranthesis. Some SNPs are labeled “untyped” because of failed assays.

TABLE 19

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
Rs1000658	211	27927511	C/T	0.79	0.80	0.591
Rs1984880	473	27927773	C/T	0.47	0.48	0.735
Rs999112	1536	27928836	C/T	0.72	0.72	0.775
Rs735880	5639	27932939	C/T	0.20	0.19	0.561
Rs2045029	17186	27944486	A/G	0.54	0.56	0.361
Rs2045028	17335	27944635	C/T			
Rs1947384	25029	27952329	C/G	0.63	0.60	0.122
Rs1947385	25111	27952411	C/T			
Rs1901744	28811	27956111	A/G	0.18	0.18	0.796
Rs1901745	28863	27956163	A/T	0.14	0.18	0.117
Rs971882	30809	27958109	A/C			
Rs1377338	40985	27968285	A/C	0.28	0.24	0.085
Rs2305452	45147	27972447	C/T	0.31	0.27	0.078
Rs2305451	45282	27972582	A/G	0.48	0.52	0.130
Rs2123472	46168	27973468	G/T	0.42	0.45	0.239
Rs2167768	46328	27973628	A/G	0.38	0.35	0.350
Rs1563055	49077	27976377	A/G			
Rs2290371	51925	27979225	C/T	0.28	0.24	0.039
Rs2290370	52141	27979441	A/G	0.85	0.84	0.551
Rs2290369	52168	27979468	C/T	0.43	0.47	0.138
Rs2874904	60852	27988152	C/T	0.26	0.23	0.132
Rs3213997	62468	27989768	A/G	0.44	0.47	0.201
Rs3213998	65572	27992872	G/T	0.83	0.80	0.223
Rs1530929	79089	28006389	A/C	0.47	0.49	0.556

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
Rs1000275	79541	28006841	C/T	0.86	0.87	0.771
Rs1000274	79790	28007090	C/T	0.54	0.56	0.510
Rs3757896	90843	28018143	A/G			
Rs3757895	90978	28018278	C/T	0.46	0.47	0.874
Rs3757894	91052	28018352	C/G	0.08	0.09	0.709
Rs3757893	91131	28018431	A/G	0.16	0.15	0.590
Rs3757892	91132	28018432	C/T	0.09	0.08	0.595
Rs3757891	94439	28021739	A/G			
Rs3757890	94621	28021921	A/T	0.98	0.96	0.167

[0240] The *ELP3* proximal SNPs were also allelotyped in the replication cohorts using the methods described herein and the primers provided in Tables 17 and 18. The replication allelotyping results for replication cohort #1 and replication cohort #2 are provided in Tables 20 and 21, respectively.

TABLE 20

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1000658	211	27927511	C/T	0.78	0.79	0.863
rs1984880	473	27927773	C/T	0.46	0.48	0.594
rs999112	1536	27928836	C/T	0.71	0.70	0.759
rs735880	5639	27932939	C/T	0.20	0.17	0.255
rs2045029	17186	27944486	A/G	0.55	0.57	0.526
rs2045028	17335	27944635	C/T			
rs1947384	25029	27952329	C/G	0.65	0.61	0.198
rs1947385	25111	27952411	C/T			
rs1901744	28811	27956111	A/G	0.19	0.18	0.674
rs1901745	28863	27956163	A/T	0.15	0.18	0.448
rs971882	30809	27958109	A/C			
rs1377338	40985	27968285	A/C	0.29	0.22	0.039
rs2305452	45147	27972447	C/T	0.31	0.26	0.067
rs2305451	45282	27972582	A/G	0.49	0.56	0.063
rs2123472	46168	27973468	G/T	0.42	0.49	0.039
rs2167768	46328	27973628	A/G	0.36	0.34	0.396
rs1563055	49077	27976377	A/G			
rs2290371	51925	27979225	C/T	0.28	0.23	0.054
rs2290370	52141	27979441	A/G	0.85	0.83	0.488
rs2290369	52168	27979468	C/T	0.41	0.49	0.036
rs2874904	60852	27988152	C/T	0.29	0.22	0.062
rs3213997	62468	27989768	A/G	0.44	0.50	0.064
rs3213998	65572	27992872	G/T	0.84	0.82	0.336
rs1530929	79089	28006389	A/C	0.48	0.52	0.311
Rs1000275	79541	28006841	C/T	0.86	0.87	0.566
Rs1000274	79790	28007090	C/T	0.54	0.59	0.159
Rs3757896	90843	28018143	A/G			
Rs3757895	90978	28018278	C/T	0.45	0.49	0.308
Rs3757894	91052	28018352	C/G	0.09	0.09	0.914
Rs3757893	91131	28018431	A/G	0.15	0.14	0.803
Rs3757892	91132	28018432	C/T	0.09	0.08	0.798
Rs3757891	94439	28021739	A/G			
Rs3757890	94621	28021921	A/T	0.98	0.95	0.159

TABLE 21

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1000658	211	27927511	C/T	0.80	0.82	0.443
rs1984880	473	27927773	C/T	0.48	0.47	0.898
rs999112	1536	27928836	C/T	0.72	0.76	0.319
rs735880	5639	27932939	C/T	0.20	0.22	0.598
rs2045029	17186	27944486	A/G	0.52	0.54	0.581
rs2045028	17335	27944635	C/T			
rs1947384	25029	27952329	C/G	0.62	0.59	0.348
rs1947385	25111	27952411	C/T			
rs1901744	28811	27956111	A/G	0.18	0.18	0.928
rs1901745	28863	27956163	A/T	0.13	0.17	0.113
rs971882	30809	27958109	A/C			
rs1377338	40985	27968285	A/C	0.27	0.27	0.961
rs2305452	45147	27972447	C/T	0.32	0.30	0.673
rs2305451	45282	27972582	A/G	0.47	0.47	0.911
rs2123472	46168	27973468	G/T	0.41	0.38	0.348
rs2167768	46328	27973628	A/G	0.39	0.37	0.664
rs1563055	49077	27976377	A/G			
rs2290371	51925	27979225	C/T	0.28	0.25	0.403
rs2290370	52141	27979441	A/G	0.85	0.84	0.939
rs2290369	52168	27979468	C/T	0.46	0.44	0.712
rs2874904	60852	27988152	C/T	0.24	0.24	0.888
rs3213997	62468	27989768	A/G	0.45	0.43	0.752
rs3213998	65572	27992872	G/T	0.81	0.78	0.373
rs1530929	79089	28006389	A/C	0.46	0.43	0.445
rs1000275	79541	28006841	C/T	0.87	0.86	0.767
rs1000274	79790	28007090	C/T	0.54	0.51	0.394
rs3757896	90843	28018143	A/G			
rs3757895	90978	28018278	C/T	0.47	0.42	0.202
rs3757894	91052	28018352	C/G	0.07	0.09	0.478
rs3757893	91131	28018431	A/G	0.17	0.16	0.653
rs3757892	91132	28018432	C/T	0.09	0.07	0.567
rs3757891	94439	28021739	A/G			
rs3757890	94621	28021921	A/T	0.97	0.97	0.728

[0241] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1B for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1B can be determined by consulting Table 19. For example, the left-most X on the left graph is at position 27927511. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0242] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to

expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than 10^{-8} were truncated at that value.

[0243] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is place at the 3' end of each gene to show the direction of transcription.

Example 6

CHDC1 Region Proximal SNPs

[0244] It has been discovered that SNP rs912428 in calponin homology (CH) domain containing 1 (*CHDC1*) is associated with occurrence of osteoarthritis in subjects.

[0245] Forty-three additional allelic variants proximal to rs912428 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 22. The chromosome positions provided in column four of Table 22 are based on Genome "Build 34" of NCBI's GenBank.

TABLE 22

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs1012628	13	243	44917643	c/t
rs1570976	13	10208	44927608	c/t
rs912436	13	15049	44932449	c/t
rs912435	13	15111	44932511	a/g
rs912433	13	15272	44932672	c/t
rs912432	13	15287	44932687	a/g
rs912431	13	15326	44932726	a/g
rs912430	13	15327	44932727	c/t
rs1408225	13	17038	44934438	c/t
rs998657	13	19391	44936791	a/g
rs1324006	13	21702	44939102	c/t
rs1924417	13	22431	44939831	c/g
rs2038728	13	22881	44940281	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs912429	13	27744	44945144	a/t
rs3742269	13	32564	44949964	a/g
rs3742270	13	32698	44950098	a/c
rs3803192	13	33104	44950504	g/t
rs3803191	13	33181	44950581	c/t
rs754106	13	33256	44950656	c/t
rs2005053	13	33543	44950943	c/t
rs1535793	13	35567	44952967	c/t
rs1886220	13	40085	44957485	c/t
rs1886219	13	40482	44957882	a/t
rs1535792	13	45641	44963041	a/t
rs1535791	13	46059	44963459	a/g
rs912428	13	48504	44965904	c/t
rs1886218	13	48919	44966319	a/c
rs1570622	13	49693	44967093	c/t
rs912427	13	49874	44967274	a/g
rs912426	13	50020	44967420	a/g
rs3068693	13	50616	44968016	-/ttt
rs1570621	13	50719	44968119	a/g
rs1886965	13	55511	44972911	c/t
rs1008849	13	65533	44982933	a/g
rs912434	13	70529	44987929	a/c
rs3889095	13	75591	44992991	c/t
rs716223	13	77266	44994666	g/t
rs2897207	13	80368	44997768	g/t
rs1570620	13	82475	44999875	a/g
rs1467605	13	92462	45009862	g/t
rs1467604	13	92480	45009880	c/t
rs1408224	13	95819	45013219	c/t
rs1408223	13	96275	45013675	c/t

Assay for Verifying and Allelotyping SNPs

[0246] The methods used to verify and allelotype the 43 proximal SNPs of Table 22 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 23 and Table 24, respectively.

TABLE 23

dbSNP rs#	Forward PCR primer	Reverse PCR primer
Rs1012628	ACGTTGGATGGATTTTCTGTGTCCCCCAAG	ACGTTGGATGTTGCAACAGAGAGAGCTCTG
Rs1570976	ACGTTGGATGTGATGTGTCTGCTGTGTTGG	ACGTTGGATGTTACATGGCGAGGTCTTAG
rs912436	ACGTTGGATGCCATATAAGGTGGTTATGGG	ACGTTGGATGCAAAACAGTTTTTCTGAGGC
rs912435	ACGTTGGATGCAAGCCAATATCCAAGACAG	ACGTTGGATGAAAAACCTGTTTGTGAGGCC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs912433	ACGTTGGATGTGCCTTCCATCCTTAACACG	ACGTTGGATGGGCTTGAGCTTAGATATGGC
rs912432	ACGTTGGATGAAATAGTTGGGTTTTGTGCC	ACGTTGGATGATTTGGTGTTAATTGCAGTG
rs912431	ACGTTGGATGTGGAAGGCACAAAACCCAAC	ACGTTGGATGCAGAAGCTAGGCTTCCTATG
rs912430	ACGTTGGATGTGGAAGGCACAAAACCCAAC	ACGTTGGATGCAGAAGCTAGGCTTCCTATG
Rs1408225	ACGTTGGATGGGGCACCATGACAATATTCC	ACGTTGGATGACACCTTGATCTTGGACTTC
rs998657	ACGTTGGATGACTGGGCCAGGGAGGAATAG	ACGTTGGATGGTTGGGGAGATAATACAGAAG
Rs1324006	ACGTTGGATGGCTGAAAACCCAAATGTGTG	ACGTTGGATGCCAGCTATCAGCTCCATTTT
Rs1924417	ACGTTGGATGACAAAAGCAAGCCTTCACAG	ACGTTGGATGGTACTGTAAAAGGTACTGTG
Rs2038728	ACGTTGGATGAAGGCTTTTGGACACAAGTC	ACGTTGGATGGCACCTCTTATGATGTTCCC
rs912429	ACGTTGGATGTTCAATTCCCCAAAGCCCTC	ACGTTGGATGGGCAAGTTCATAACCTCTC
Rs3742269	ACGTTGGATGGAGAAAAGAGAACGAGAAGG	ACGTTGGATGTAAATGACAGCAGTCTGGAG
Rs3742270	ACGTTGGATGCTAAAACCAAAGCTGACGGG	ACGTTGGATGTTCTGCTCCTGTGGCATAGC
Rs3803192	ACGTTGGATGTCCTTTTGCTTCTGCGATGC	ACGTTGGATGTGCTTCCCACATCAGTTCTTG
Rs3803191	ACGTTGGATGCTGTCTGTACATTACCAGGC	ACGTTGGATGAATAGCAGCTGGAGGATCTC
rs754106	ACGTTGGATGTTCTTACCATCCAGCAAGGC	ACGTTGGATGGCCTGGTAATGTACAGACAG
Rs2005053	ACGTTGGATGCTGTTGCTAGCTTGGATTTG	ACGTTGGATGTTCCCTGTCTTTCTGGCAT
Rs1535793	ACGTTGGATGAACAAAGAGGAACAGAGCCC	ACGTTGGATGGCATAAGCCCCCTTTCTAG
Rs1886220	ACGTTGGATGTCACCGTGTAGCGAGAATG	ACGTTGGATGTAAATCCAGCACTTTGGGAG
Rs1886219	ACGTTGGATGTGTAAGTGGATTGCTGGAG	ACGTTGGATGTACATCAATAGCCGAGGAAG
Rs1535792	ACGTTGGATGCTGTATATCAGTGAAGTCTC	ACGTTGGATGCAGAGAAGAAACATCTCAGC
Rs1535791	ACGTTGGATGGAGGGTTTATCCTTACAATTG	ACGTTGGATGTTTATAGGGTCCCTTGATAAG
rs912428	ACGTTGGATGACTACATCCATTCCAGGGAG	ACGTTGGATGTCAGATCAGAGTGAGTTAG
Rs1886218	ACGTTGGATGTCCCGAAAACAAGTCAAGAC	ACGTTGGATGAGTCCAGGCAAAACAGTAAG
Rs1570622	ACGTTGGATGATAGCTGCCACACTCTTTAG	ACGTTGGATGGCGCAGTTTAGAAAAACCTG
rs912427	ACGTTGGATGTAGGGTTCTCGATGGGTATG	ACGTTGGATGTTTGCCCTGGTCACTTTAGG
rs912426	ACGTTGGATGTTAGAGGATGCATAGGCCAG	ACGTTGGATGAAGTCACTTACTGCATGGTC
Rs3068693	ACGTTGGATGAAATTGGCCACATGGAATCC	ACGTTGGATGCTACCTTTAACATCCCTGTC
Rs1570621	ACGTTGGATGAATTAAGAATGGCAGCTATG	ACGTTGGATGGTTTAAACTAAAAACAC
Rs1886965	ACGTTGGATGCTGCTAAGGATATGTGTTTCC	ACGTTGGATGACACCAGTGCTCAGTATTTG
Rs1008849	ACGTTGGATGGCAGTTGTGAATTGTGCAGC	ACGTTGGATGTGGTGCAGAACATGTCAGAC
rs912434	ACGTTGGATGTTCTGACATGTACAGACGTG	ACGTTGGATGTCCTGGGAAATCTTCCATC
Rs3889095	ACGTTGGATGAAGGTAATGATATGTCCCC	ACGTTGGATGCGCATTTTACAGAGACATTG
rs716223	ACGTTGGATGACACTGTCTCTAGAAGCAGG	ACGTTGGATGGAAGCAGGAAAAGAGTGAGG
Rs2897207	ACGTTGGATGTCAGCCTCCAGAACTATGAG	ACGTTGGATGAACAGAGAGAGACCCTGTCT
Rs1570620	ACGTTGGATGCTGTTCTGCTTGTATATGG	ACGTTGGATGGAAGGAAGTCTATTAGCCC
Rs1467605	ACGTTGGATGATGTTACAGGGTGGTAAGCG	ACGTTGGATGTAAAGTTGCCACGCTTCTCC
Rs1467604	ACGTTGGATGATATACGGCATGTTACAGGG	ACGTTGGATGTTAAAGTTGCCACGCTTCTC
Rs1408224	ACGTTGGATGACTTCCCACTCCTCTAGACA	ACGTTGGATGTATTGGCTGGGTAGCACTCC
Rs1408223	ACGTTGGATGTCATTACCAGTTCCACAGAG	ACGTTGGATGTTGAGACATCATGAGGAGTG

TABLE 24

dbSNP rs#	Extend Primer	Term Mix
Rs1012628	CTGTGTCCCCCAAGTCTTTG	ACG
Rs1570976	TTGGCATTTCCTTGAGAA	ACT
rs912436	AGGTGGTTATGGGTTTGTCACCTCA	ACT
rs912435	TCCAAAAAGCCCAAGAAATTCT	ACT
rs912433	CCTTAACACGTTTATAATAGATTA	ACG
rs912432	GTGCCTTCCATCCTTAACAC	ACT
rs912431	GGCACAAAACCCAACTATTTTTTC	ACG
rs912430	GCACAAAACCCAACTATTTTTTCC	ACT
Rs1408225	CCTCAGACTGGGTGGCTTA	ACT
rs998657	CACCCACCTGAGGGAGGC	ACT
Rs1324006	GATACCTTGAAGAATTTTTAAAAC	ACG
Rs1924417	TTTAGGCACATTTGTACTTATAAA	ACT
Rs2038728	TGGACACAAGTCCATGCAACA	ACG
rs912429	CTGTGACAGGTGCTATTATCA	CGT
Rs3742269	TTTTGGACCGATTTCGGTG	ACT
Rs3742270	GCTGACGGGGATTCCCTTTA	ACT
Rs3803192	GATGCACTAAAAGCAGCAATGT	ACT
Rs3803191	TCCAGCCTTCATATTTTCCTC	ACG
rs754106	ATCCAGCAAGGCACTTAGAAT	ACT
Rs2005053	TGTGGCCTTCAGATGCTTACAT	ACG
Rs1535793	GAGGAACAGAGCCCAAAGGACA	ACT
Rs1886220	CTGACCTCGTGATCCGCC	ACG
Rs1886219	ACTGGATTTGCTGGAGTTAAGAA	CGT
Rs1535792	TATCAGTGAAGTGCCTTTCTTTT	CGT
Rs1535791	TTATCCTTACAATTGAAGAAAGGA	ACT
rs912428	CCATTCCAGGGAGACTCCCA	ACT
Rs1886218	GAAAACAAGTCAAGACATTTATTG	ACT
Rs1570622	CTGCCACACTCTTTAGATGAAGTT	ACG
rs912427	GGGAGATGACAGAACAAAAC	ACT
rs912426	AGGTGCCAAGTGTAGAAAGAAAC	ACG
Rs3068693	GCCTCACATTGTTTTTTTTTTTTT	ACT
Rs1570621	TCGGTCATAACTTTAATGAAGG	ACG
Rs1886965	TGATTTTATGACTCACATTATTC	ACT
Rs1008849	GTGAATTGTGCAGCTATAAACATG	ACG
rs912434	AGACGTGCCAGCTATGATA	ACT
Rs3889095	TCCCCATAACATTTTCAGCAT	ACT
rs716223	GTGGTTTGTATTTCCAGTGTCA	ACT
Rs2897207	AACTATGAGAAATAAATGTGTGGG	ACT
Rs1570620	TTGATATGGTTCTTGGTTGTTGG	ACG
Rs1467605	GTAAGCGCTAGAAAGAAAAATAA	ACT
Rs1467604	ACGGCATGTTACAGGGTGGTAAG	ACG
Rs1408224	GGGCACACATTCAGAACTGCCC	ACG
Rs1408223	ACAGAGGAAGACCAAATGACA	ACG

Genetic Analysis

[0247] Allelotyping results from the discovery cohort are shown for cases and controls in Table 25. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where “AF” is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs1570976 has the following case and control allele frequencies: case A1 (C) = 0.49; case A2 (T) = 0.51; control A1 (C) = 0.53; and control A2 (T) = 0.47, where the nucleotide is provided in paranthesis. Some SNPs are labeled “untyped” because of failed assays.

TABLE 25

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1012628	243	44917643	C/T	0.70	0.70	0.768
rs1570976	10208	44927608	C/T	0.51	0.47	0.125
rs912436	15049	44932449	C/T	0.98	untyped	
rs912435	15111	44932511	A/G	0.64	0.36	~0.0001
rs912433	15272	44932672	C/T	0.22	0.23	0.581
rs912432	15287	44932687	A/G	0.46	0.44	0.282
rs912431	15326	44932726	A/G	0.46	0.46	0.969
rs912430	15327	44932727	C/T	0.20	0.19	0.584
rs1408225	17038	44934438	C/T			
rs998657	19391	44936791	A/G	0.47	0.44	0.254
rs1324006	21702	44939102	C/T	0.55	0.53	0.419
rs1924417	22431	44939831	C/G	0.53	0.49	0.108
rs2038728	22881	44940281	A/G	0.34	0.38	0.082
rs912429	27744	44945144	A/T			
rs3742269	32564	44949964	A/G	0.83	0.83	0.967
rs3742270	32698	44950098	A/C	0.53	0.50	0.170
rs3803192	33104	44950504	G/T			
rs3803191	33181	44950581	C/T			
rs754106	33256	44950656	C/T	0.40	0.41	0.714
rs2005053	33543	44950943	C/T	0.40	0.40	0.877
rs1535793	35567	44952967	C/T	0.26	0.26	0.910
rs1886220	40085	44957485	C/T			
rs1886219	40482	44957882	A/T	0.21	0.22	0.867
rs1535792	45641	44963041	A/T	0.73	0.71	0.550
rs1535791	46059	44963459	A/G	0.08	0.15	0.009
rs912428	48504	44965904	C/T			
rs1886218	48919	44966319	A/C			
rs1570622	49693	44967093	C/T	0.73	0.75	0.451
rs912427	49874	44967274	A/G	0.68	0.70	0.352
rs912426	50020	44967420	A/G	0.76	0.77	0.680
rs3068693	50616	44968016	-/TTT	0.22	0.21	0.597
rs1570621	50719	44968119	A/G	0.19	0.18	0.569
rs1886965	55511	44972911	C/T			
rs1008849	65533	44982933	A/G	0.48	0.43	0.160
rs912434	70529	44987929	A/C	0.23	0.23	0.988
rs3889095	75591	44992991	C/T	0.90	0.90	0.880
rs716223	77266	44994666	G/T	0.91	0.90	0.981
rs2897207	80368	44997768	G/T	0.46	0.46	0.921
rs1570620	82475	44999875	A/G	0.67	0.68	0.738
rs1467605	92462	45009862	G/T	0.29	0.22	0.044

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1467604	92480	45009880	C/T	0.68	0.67	0.537
rs1408224	95819	45013219	C/T	0.66	0.65	0.683
rs1408223	96275	45013675	C/T	0.29	0.28	0.587

[0248] The *CHDC1* proximal SNPs were also allelotyped in the replication cohorts using the methods described herein and the primers provided in Tables 23 and 24. The replication allelotyping results for replication cohort #1 and replication cohort #2 are provided in Tables 26 and 27, respectively.

TABLE 26

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1012628	243	44917643	C/T	0.69	0.72	0.337
rs1570976	10208	44927608	C/T	0.48	0.46	0.490
rs912436	15049	44932449	C/T			
rs912435	15111	44932511	A/G	0.16	untyped	0.637
rs912433	15272	44932672	C/T	0.28	0.28	0.984
rs912432	15287	44932687	A/G	0.46	0.42	0.260
rs912431	15326	44932726	A/G	0.46	0.48	0.602
rs912430	15327	44932727	C/T	0.18	0.20	0.476
rs1408225	17038	44934438	C/T			
rs998657	19391	44936791	A/G	0.46	0.43	0.380
rs1324006	21702	44939102	C/T	0.54	0.53	0.811
rs1924417	22431	44939831	C/G	0.51	0.49	0.440
rs2038728	22881	44940281	A/G	0.35	0.39	0.181
rs912429	27744	44945144	A/T			
rs3742269	32564	44949964	A/G	0.84	0.85	0.911
rs3742270	32698	44950098	A/C	0.56	0.50	0.090
rs3803192	33104	44950504	G/T			
rs3803191	33181	44950581	C/T			
rs754106	33256	44950656	C/T	0.40	0.40	0.827
rs2005053	33543	44950943	C/T	0.40	0.37	0.328
rs1535793	35567	44952967	C/T	0.27	0.24	0.259
rs1886220	40085	44957485	C/T			
rs1886219	40482	44957882	A/T	0.22	0.19	0.302
rs1535792	45641	44963041	A/T	0.73	0.76	0.435
rs1535791	46059	44963459	A/G	0.08	0.08	0.958
rs912428	48504	44965904	C/T			
rs1886218	48919	44966319	A/C			
rs1570622	49693	44967093	C/T	0.71	0.79	0.007
rs912427	49874	44967274	A/G	0.65	0.73	0.007
rs912426	50020	44967420	A/G	0.74	0.80	0.047
rs3068693	50616	44968016	-/TTT	0.25	0.21	0.236
rs1570621	50719	44968119	A/G	0.22	0.15	0.028
rs1886965	55511	44972911	C/T			
rs1008849	65533	44982933	A/G	0.47	untyped	NA
rs912434	70529	44987929	A/C	0.24	0.19	0.083
rs3889095	75591	44992991	C/T	0.91	0.91	0.867
rs716223	77266	44994666	G/T	0.91	0.93	0.598
rs2897207	80368	44997768	G/T	0.48	0.45	0.321
rs1570620	82475	44999875	A/G	0.66	0.72	0.034
rs1467605	92462	45009862	G/T	0.29	0.22	0.044
rs1467604	92480	45009880	C/T	0.66	0.70	0.307

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1408224	95819	45013219	C/T	0.64	0.67	0.312
rs1408223	96275	45013675	C/T	0.31	0.23	0.028

TABLE 27

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1012628	243	44917643	C/T	0.71	0.68	0.438
rs1570976	10208	44927608	C/T	0.55	0.50	0.159
rs912436	15049	44932449	C/T			
rs912435	15111	44932511	A/G	0.66	untyped	
rs912433	15272	44932672	C/T	0.14	0.17	0.479
rs912432	15287	44932687	A/G	0.47	0.46	0.806
rs912431	15326	44932726	A/G	0.46	0.44	0.513
rs912430	15327	44932727	C/T	0.23	0.17	0.084
rs1408225	17038	44934438	C/T			
rs998657	19391	44936791	A/G	0.48	0.45	0.518
rs1324006	21702	44939102	C/T	0.55	0.52	0.324
rs1924417	22431	44939831	C/G	0.54	0.49	0.123
rs2038728	22881	44940281	A/G	0.34	0.37	0.295
rs912429	27744	44945144	A/T			
rs3742269	32564	44949964	A/G	0.82	0.82	0.861
rs3742270	32698	44950098	A/C	0.50	0.49	0.873
rs3803192	33104	44950504	G/T			
rs3803191	33181	44950581	C/T			
rs754106	33256	44950656	C/T	0.41	0.44	0.346
rs2005053	33543	44950943	C/T	0.40	0.44	0.302
rs1535793	35567	44952967	C/T	0.25	0.31	0.096
rs1886220	40085	44957485	C/T			
rs1886219	40482	44957882	A/T	0.20	0.27	0.053
rs1535792	45641	44963041	A/T	0.73	0.63	0.007
rs1535791	46059	44963459	A/G	NA	0.27	NA
rs912428	48504	44965904	C/T			
rs1886218	48919	44966319	A/C			
rs1570622	49693	44967093	C/T	0.75	0.67	0.040
rs912427	49874	44967274	A/G	0.71	0.64	0.059
rs912426	50020	44967420	A/G	0.78	0.72	0.065
rs3068693	50616	44968016	-/TTT	0.19	0.21	0.520
rs1570621	50719	44968119	A/G	0.15	0.21	0.077
rs1886965	55511	44972911	C/T			
rs1008849	65533	44982933	A/G	0.49	0.43	0.138
rs912434	70529	44987929	A/C	0.21	0.28	0.027
rs3889095	75591	44992991	C/T	0.89	0.88	0.583
rs716223	77266	44994666	G/T	0.90	0.87	0.368
rs2897207	80368	44997768	G/T	0.44	0.48	0.276
rs1570620	82475	44999875	A/G	0.70	0.62	0.026
rs1467605	92462	45009862	G/T			
rs1467604	92480	45009880	C/T	0.71	0.62	0.018
rs1408224	95819	45013219	C/T	0.68	0.61	0.060
rs1408223	96275	45013675	C/T	0.27	0.34	0.023

[0249] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values

were plotted in Figure 1C for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1C can be determined by consulting Table 25. For example, the left-most X on the left graph is at position 44917643. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0250] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than 10^{-8} were truncated at that value.

[0251] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

Example 7

ERG Region Proximal SNPs

[0252] It has been discovered that SNP rs1888475 in v-ets erythroblastosis virus E26 oncogene like (*ERG*) is associated with occurrence of osteoarthritis in subjects. One hundred sixty-six additional allelic variants proximal to rs1888475 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 28. The chromosome positions provided in column four of Table 28 are based on Genome "Build 34" of NCBI's GenBank.

TABLE 28

dbSNP rs#	Chromosome	Position in SEQ ID NO: 4	Chromosome Position	Allele Variants
rs2898353	21	231	38783681	a/t
rs960818	21	882	38784332	a/g
rs960819	21	960	38784410	a/c
rs2410034	21	1194	38784644	a/c
rs2836437	21	1530	38784980	a/g
rs2836438	21	1673	38785123	a/g
rs2836439	21	2096	38785546	c/t
rs2836440	21	2285	38785735	a/g
rs2226683	21	5873	38789323	c/t
rs2836441	21	7256	38790706	a/g
rs2836442	21	7988	38791438	a/g
rs2836443	21	8222	38791672	g/t
rs2836444	21	8381	38791831	c/t
rs3787906	21	8814	38792264	c/t
rs3838108	21	8915	38792365	-/c
rs2836445	21	9642	38793092	a/g
rs2836446	21	9902	38793352	a/t
rs3787908	21	10619	38794069	a/g
rs2836447	21	10927	38794377	c/t
rs2836448	21	11032	38794482	c/t
rs2836450	21	14377	38797827	c/t
rs2836451	21	15608	38799058	c/t
rs1015022	21	15928	38799378	c/g
rs2836452	21	16296	38799746	a/g
rs2836453	21	17598	38801048	a/t
rs3787909	21	19272	38802722	a/g
rs2836454	21	20084	38803534	a/g
rs2836455	21	20577	38804027	a/t
rs2155718	21	28051	38811501	a/g
rs2836456	21	29466	38812916	a/g
rs2836457	21	29530	38812980	c/t
rs2836458	21	29987	38813437	a/g
rs2032323	21	30012	38813462	c/t
rs2051400	21	30322	38813772	g/t
rs2836459	21	32216	38815666	c/t
rs2836460	21	32516	38815966	c/t
rs2836461	21	32544	38815994	a/g
rs2836462	21	32746	38816196	a/g
rs2836463	21	33137	38816587	g/t
rs2836464	21	33538	38816988	a/g
rs2836465	21	33798	38817248	c/t
rs2836466	21	33802	38817252	a/c
rs2836467	21	33964	38817414	c/t
rs3827204	21	34132	38817582	a/g
rs2836468	21	34210	38817660	c/t
rs3787911	21	34317	38817767	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 4	Chromosome Position	Allele Variants
rs2836469	21	34499	38817949	c/t
rs2836470	21	34753	38818203	a/c
rs2212599	21	34845	38818295	c/t
rs2836472	21	35335	38818785	c/t
rs2836473	21	36423	38819873	c/t
rs1888469	21	36450	38819900	a/g
rs1888470	21	36481	38819931	g/t
rs2032322	21	38447	38821897	c/g
rs2410035	21	38784	38822234	c/t
rs1573332	21	39387	38822837	a/t
rs2836474	21	39458	38822908	c/t
rs2836475	21	39822	38823272	c/g
rs3787914	21	40305	38823755	c/g
rs1888471	21	40869	38824319	c/t
rs1888472	21	40926	38824376	c/t
rs1888473	21	41010	38824460	c/t
rs1888474	21	41134	38824584	c/t
rs2836476	21	41984	38825434	a/g
rs3787916	21	42172	38825622	a/t
rs2836477	21	42753	38826203	g/t
rs970043	21	43011	38826461	c/t
rs2212600	21	43176	38826626	a/g
rs2836478	21	43320	38826770	g/t
rs2836479	21	43381	38826831	a/t
rs1475877	21	44142	38827592	a/g
rs2836480	21	44383	38827833	a/g
rs2836481	21	44726	38828176	c/t
rs2836483	21	45087	38828537	a/g
rs2836484	21	45141	38828591	c/t
rs2836485	21	45359	38828809	c/g
rs2836486	21	45421	38828871	c/t
rs2836487	21	45456	38828906	c/t
rs1893199	21	45467	38828917	c/t
rs2836488	21	45486	38828936	c/t
rs1893200	21	45709	38829159	a/g
rs1893201	21	45716	38829166	a/g
rs2836489	21	47626	38831076	c/t
rs1888475	21	49413	38832863	a/g
rs2836490	21	49796	38833246	c/t
rs2836491	21	49962	38833412	a/g
rs2836492	21	50075	38833525	c/t
rs2836493	21	50093	38833543	a/g
rs2836494	21	50571	38834021	c/t
rs2836495	21	50615	38834065	a/g
rs2898354	21	50780	38834230	a/g
rs3065390	21	50851	38834301	-/ta
rs2836496	21	51459	38834909	a/c

dbSNP rs#	Chromosome	Position in SEQ ID NO: 4	Chromosome Position	Allele Variants
rs2836497	21	53193	38836643	c/t
rs2836498	21	53702	38837152	c/t
rs2836499	21	53736	38837186	a/c
rs2836500	21	53795	38837245	c/t
rs2836501	21	54109	38837559	a/t
rs2836502	21	54126	38837576	c/t
rs2836503	21	54230	38837680	a/c
rs2836504	21	54894	38838344	c/t
rs3787917	21	55455	38838905	a/g
rs2836505	21	55499	38838949	a/g
rs2836506	21	56522	38839972	c/t
rs2836507	21	56662	38840112	c/t
rs2836508	21	56954	38840404	a/g
rs2836509	21	57267	38840717	a/g
rs2836510	21	58282	38841732	a/g
rs2836511	21	58916	38842366	a/c
rs2212601	21	59544	38842994	c/g
rs2212602	21	59666	38843116	c/t
rs2226682	21	59913	38843363	a/t
rs2836512	21	66846	38850296	a/g
rs2836513	21	67245	38850695	g/t
rs1999328	21	67652	38851102	a/c
rs2212603	21	67955	38851405	a/g
rs3787919	21	67966	38851416	a/c
rs2836514	21	68420	38851870	a/g
rs1023153	21	70226	38853676	a/g
rs1023372	21	70810	38854260	c/t
rs2212604	21	72246	38855696	a/g
rs2226684	21	73330	38856780	g/t
rs2212605	21	73457	38856907	c/t
rs2187307	21	74389	38857839	a/g
rs3065412	21	74638	38858088	-/aa
rs2898355	21	74640	38858090	a/c
rs2836518	21	75358	38858808	a/c
rs3838110	21	75952	38859402	-/g
rs2836519	21	76098	38859548	a/g
rs3827207	21	77836	38861286	a/g
rs2836520	21	78449	38861899	a/c
rs2836521	21	78507	38861957	g/t
rs2836522	21	80031	38863481	g/t
rs2836523	21	81695	38865145	c/t
rs2836524	21	82775	38866225	a/g
rs2836525	21	82795	38866245	a/g
rs3833350	21	84611	38868061	-/c
rs2836526	21	84657	38868107	c/t
rs2836527	21	84693	38868143	a/c
rs3834676	21	85020	38868470	-/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 4	Chromosome Position	Allele Variants
rs2836528	21	85048	38868498	c/t
rs3761364	21	85100	38868550	c/t
rs2836529	21	85325	38868775	a/c
rs2836530	21	85452	38868902	c/t
rs3761366	21	85868	38869318	a/g
rs2836531	21	85936	38869386	a/g
rs2836532	21	85990	38869440	a/t
rs2836533	21	86139	38869589	c/t
rs2836534	21	86497	38869947	c/t
rs2836535	21	87236	38870686	a/g
rs2836536	21	87248	38870698	c/t
rs3827208	21	87533	38870983	c/g
rs715860	21	87912	38871362	a/g
rs717231	21	88108	38871558	g/t
rs2836537	21	88494	38871944	a/c
rs2836538	21	89598	38873048	a/c
rs2836539	21	90235	38873685	a/t
rs2836540	21	91287	38874737	g/t
rs2836541	21	91359	38874809	c/t
rs2836542	21	92384	38875834	a/c
rs2836543	21	92410	38875860	c/t
rs881837	21	92900	38876350	c/t
rs3949052	21	94495	38877945	a/g
rs2065307	21	94512	38877962	a/g
rs3216105	21	97777	38881227	-/a
rs2073427	21	98333	38881783	c/t

Assay for Verifying and Allelotyping SNPs

[0253] The methods used to verify and allelotype the 166 proximal SNPs of Table 28 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 29 and Table 30, respectively.

TABLE 29

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2898353	ACGTTGGATGAATGTGAATGTGGAGGTAGC	ACGTTGGATGCTCCCTTGCTGGTTTTTTTG
rs960818	ACGTTGGATGTGGGATTTTTCCCAGAAGAG	ACGTTGGATGCTGTGCAGAGAAACATGATG
rs960819	ACGTTGGATGCTGTCTCCCTTCTCTTTATC	ACGTTGGATGCATCATGTTTCTCTGCACAG
rs2410034	ACGTTGGATGTTTAGAGACATTTCTCCTAG	ACGTTGGATGTTAGGATGATGTTAGTTTGG
rs2836437	ACGTTGGATGAGCTTCTGCGATATCAGTGG	ACGTTGGATGTTCTGTGCAGCACATTCTCC
rs2836438	ACGTTGGATGAACATGTCTTGCCAAGCTC	ACGTTGGATGCCACTGTGACCTCTGGATTT
rs2836439	ACGTTGGATGCCTAGTGTATAAAGTGATGC	ACGTTGGATGTCCTTTCTAGGCACCAATAC
rs2836440	ACGTTGGATGAGATCCTAACCAACCACAGC	ACGTTGGATGAGGTAGGTAGATACAAGGCC
rs2226683	ACGTTGGATGAATATGGCTCCTATAGACAG	ACGTTGGATGTTTTGGGTCACAAAATCAAG

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2836441	ACGTTGGATGTTACCTTAATAGTGCTGGCC	ACGTTGGATGACTTTCTGGTCAGAGAGAAG
rs2836442	ACGTTGGATGCAAGGACTCTAGGCTTACAG	ACGTTGGATGGGACATTTGTAGTCACTTC
rs2836443	ACGTTGGATGGGGCCCCATTACATGTCTAA	ACGTTGGATGTTCTGCTGTACTTCTTCGAG
rs2836444	ACGTTGGATGCTGCAACCAGGAATTGTCAG	ACGTTGGATGAGGACCATAAAGAGGTGTG
rs3787906	ACGTTGGATGTGAAAAGAGCGGAAATCAAC	ACGTTGGATGGTAAGAAAATCATTCTGTGG
rs3838108	ACGTTGGATGATGAATAAGATGGCAGGCTG	ACGTTGGATGAAGCTGCCAGATAAAACAG
rs2836445	ACGTTGGATGCATTTCCAAAATTAGACGCAG	ACGTTGGATGAAAAAGAGAAAAACAGATGC
rs2836446	ACGTTGGATGGTGCCTTGTCTATCAAGAG	ACGTTGGATGAGCATCCAAGCCTGGTAATC
rs3787908	ACGTTGGATGAATCACCACACTAGACCAGC	ACGTTGGATGCATGCAAGGGAAATGTGTGC
rs2836447	ACGTTGGATGATCTCCTCTCTTGTCTGCTG	ACGTTGGATGGAGGAAGGTTAGGAGCTAAG
rs2836448	ACGTTGGATGTGTAGGGATGTATAGGGCAG	ACGTTGGATGAAAGAGAGGAGATCCGCTCG
rs2836450	ACGTTGGATGTGTGGGCATCAGATGACAAC	ACGTTGGATGATCCCGTTAAATGCACCGAC
rs2836451	ACGTTGGATGCAGACAAACAACGTCAACC	ACGTTGGATGGTATTTCTTTTCTCGCCGC
rs1015022	ACGTTGGATGTCGAGCCAGCGCTTTTATC	ACGTTGGATGGTAACAGTCGTACATTCCGG
rs2836452	ACGTTGGATGATCACTGACACAGTCATGAG	ACGTTGGATGCCAGTAACCTTTCAGGTTTG
rs2836453	ACGTTGGATGTGTATTTCCCAAGATGGCCC	ACGTTGGATGCCTCACTTTCTGATGGAAGC
rs3787909	ACGTTGGATGACTTCTCAGTGTTCTGGCTG	ACGTTGGATGCGTCACTCTCTGTTTCATGG
rs2836454	ACGTTGGATGAGGAATGATTACAAACCTCC	ACGTTGGATGGAATGTTCAAATGTAGGGTGG
rs2836455	ACGTTGGATGGGTCTATTGCTGTGACATT	ACGTTGGATGCATCCCAATTTTAAGCAAG
rs2155718	ACGTTGGATGAGAACTCTCACACACAGCTG	ACGTTGGATGTGCCTCTTATTACAGCCCTG
rs2836456	ACGTTGGATGGGGATTGTCTGATCTCCTTG	ACGTTGGATGCCAGCTTTCCTTGTGCATG
rs2836457	ACGTTGGATGAACTCCTGGAATGAGTCACC	ACGTTGGATGATGCACAAAGGAAAGCTGGG
rs2836458	ACGTTGGATGATCACTTAGAAGCCCAGCAG	ACGTTGGATGTGATGCACACTCACTGAAGC
rs2032323	ACGTTGGATGGTAGCCGCACTTTGAGATGC	ACGTTGGATGAGCACAGAGTCGAGGAGGAG
rs2051400	ACGTTGGATGACAGACCTCAGACCAAAGTC	ACGTTGGATGTTTGTCTAGAGTAACCCCC
rs2836459	ACGTTGGATGGCAAGAATGTTACTTTCTGG	ACGTTGGATGCCATCAAATAGTTGGTTGTC
rs2836460	ACGTTGGATGCAATATCTGAGTTTACCCC	ACGTTGGATGGTAGATGAGAATCCGTGTG
rs2836461	ACGTTGGATGGTTACCCACACGGAATTCTC	ACGTTGGATGCCAGATCCAGGTTCTTTCTG
rs2836462	ACGTTGGATGTCTCCTCCGTATGTCTCCAT	ACGTTGGATGATCCCGGAACTCTCTGTTTC
rs2836463	ACGTTGGATGGCACTATTTGACTTGAGCTC	ACGTTGGATGAATTCAAGCCAGAAAGGCTC
rs2836464	ACGTTGGATGGTCTTTTTACCCCCAGTAAAG	ACGTTGGATGATAAGCAAAGGACCTTTGG
rs2836465	ACGTTGGATGTGAGCTCTTGTGTTTTGCC	ACGTTGGATGGAGAATTCTCCAGCCTTCTC
rs2836466	ACGTTGGATGTGAGCTCTTGTGTTTTGCC	ACGTTGGATGGAGAATTCTCCAGCCTTCTC
rs2836467	ACGTTGGATGGACTCTGCTCATTTCTTTGG	ACGTTGGATGAAGAGTAGGGGTAGATGCAG
rs3827204	ACGTTGGATGTGAAGATCACACGTGGTGTA	ACGTTGGATGGGGTGAATGCCAAAAAGAGG
rs2836468	ACGTTGGATGTAGAGGCAGGAAAGAGCATG	ACGTTGGATGTTTTTGGCATTACCCCTCTC
rs3787911	ACGTTGGATGTAACCTCTTCTGGATTCTGG	ACGTTGGATGTCATGTGCTCTGAGAGCATC
rs2836469	ACGTTGGATGATTTCTCTACCTCATCCCC	ACGTTGGATGGGTTGAAGTCACGTAACAGC
rs2836470	ACGTTGGATGCCACTGTTAATCGTATTGCC	ACGTTGGATGACGGACTGAAAGCCAAATGG
rs2212599	ACGTTGGATGAGGAGTTATTCTTCCCAAC	ACGTTGGATGCAGTGGTCCATTAAGAATCC
rs2836472	ACGTTGGATGGAGTATCGTTCTCTATCATG	ACGTTGGATGTAAAAGAGTCAGAGCAGGAC
rs2836473	ACGTTGGATGTCTCAGCCAGAGTTTTGACC	ACGTTGGATGAATCAACGCCTCCTCTTCAG
rs1888469	ACGTTGGATGACCACCAGGAAGGGTCTGAA	ACGTTGGATGGAGGATCAGAGGCAGAAAAC
rs1888470	ACGTTGGATGGCGTTGATTGCAGTTTTCTG	ACGTTGGATGTTCTTTGGCCTCCGTGTAAG
rs2032322	ACGTTGGATGTGATACTCTGTTGAGCCTCC	ACGTTGGATGGGGGAGCAGTGATGAGTTAT
rs2410035	ACGTTGGATGAATCACTTGAACCCAGGAGG	ACGTTGGATGTTTTGAGACGGAGTTTCGC
rs1573332	ACGTTGGATGGGGTGAACTTTACAGAGAGG	ACGTTGGATGCTGCCAGACAGTTTTGAGAC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2836474	ACGTTGGATGAATTCTGCACAGGAGAGTCC	ACGTTGGATGCAGGAAATGAAGATGTCGCC
rs2836475	ACGTTGGATGAGTTCTACATGGGAAGCTGC	ACGTTGGATGATATCTGTGTCTACAGGCC
rs3787914	ACGTTGGATGGGCTGAAGGCTAAAATCACC	ACGTTGGATGGTCTGAGAAGTAGGAATGGC
rs1888471	ACGTTGGATGACTGAGGCAATTGTGTAGAC	ACGTTGGATGTTGACTTTGTTTTGAGAGGC
rs1888472	ACGTTGGATGTTGCCTCTCAAAACAAAGTC	ACGTTGGATGCTATTATTCTGGAAGCAGCC
rs1888473	ACGTTGGATGAGAAAGTTCAGTTCTCAGCC	ACGTTGGATGTGTTTGCTCCTGTGAGTAAC
rs1888474	ACGTTGGATGTGTTATGTGAGTCCAGGGTG	ACGTTGGATGTCTTGTATGTGGGTGGGTG
rs2836476	ACGTTGGATGTTACCTGTGACCTCATTTGG	ACGTTGGATGGAACACACAACATACGGTAC
rs3787916	ACGTTGGATGAAGGCATCTCAGTCATTCTC	ACGTTGGATGTGAGTTTGACACAAAGAAGC
rs2836477	ACGTTGGATGTTTAGCTCTCCTGGATGATG	ACGTTGGATGCCATGATTAGTGCATGAAGG
rs970043	ACGTTGGATGTATAACTCCCCTCTCTCCTG	ACGTTGGATGAGAGCAGACCCCTTATCAGAG
rs2212600	ACGTTGGATGGAAACAGGTGTTCAATTTGGC	ACGTTGGATGTCTGCATGAACCAGTAAGTC
rs2836478	ACGTTGGATGAGCTATTGAGTGTCACTTGC	ACGTTGGATGCAGAAGCTTCTGACTTCAAC
rs2836479	ACGTTGGATGAGTAGCCATCCTAATAGGTG	ACGTTGGATGAGCAAGTGACACTCAATAGC
rs1475877	ACGTTGGATGAATCAACACTCCCCGTGTTTC	ACGTTGGATGGGTACCTAGAGTAGTCCAAG
rs2836480	ACGTTGGATGTACCAAACCCACTGTACATC	ACGTTGGATGCATAACCTAACACATTGTGGG
rs2836481	ACGTTGGATGTAAGAAGTCTTTCTCCCCC	ACGTTGGATGGCTGCTTCTTTTATAAGAGG
rs2836483	ACGTTGGATGCACTGAGGTAATCTCCAACC	ACGTTGGATGGGTGGAGATATGGCTTGATG
rs2836484	ACGTTGGATGAAGCCCACCAGAGTCATCAA	ACGTTGGATGACTACTGACCAGCTTTCCAG
rs2836485	ACGTTGGATGTTCTAAGTGAAGCCCTCCTC	ACGTTGGATGTACAGCTGTGCAAACAGTTG
rs2836486	ACGTTGGATGCATGGTCTGTTGCCTCTAAG	ACGTTGGATGCCCTAGCATTTTATGCATCC
rs2836487	ACGTTGGATGTGAATACCCACTAGGTCTCG	ACGTTGGATGCCACCCTAACTTAGAGGC
rs1893199	ACGTTGGATGGGCAACAGACCATGGTTTTG	ACGTTGGATGCTTCCCTTCAACATGCACTG
rs2836488	ACGTTGGATGGGCAACAGACCATGGTTTTG	ACGTTGGATGCTTCCCTTCAACATGCACTG
rs1893200	ACGTTGGATGAGTTAAGTCTTCGCATAACC	ACGTTGGATGCCTCTCACACACTAAATCTTG
rs1893201	ACGTTGGATGGTCTTCGCATAACCAAAACAG	ACGTTGGATGCCTCTCACACACTAAATCTTG
rs2836489	ACGTTGGATGGTCAACCATGGAGCTTGAAC	ACGTTGGATGAGAAGACATGTGGGCTTGTG
rs1888475	ACGTTGGATGACCCCTGGCAAGTGAATTAC	ACGTTGGATGGGGAGGTGGATGTTCTTATC
rs2836490	ACGTTGGATGAAAGGCAGAGCTAAAGCAAG	ACGTTGGATGAGCACAACCCAGCAATGCAG
rs2836491	ACGTTGGATGACAACCTTGGAGTGGAAGGG	ACGTTGGATGATCCAGATGGATTCCACAGC
rs2836492	ACGTTGGATGACATATGGGCATGGAAGAGC	ACGTTGGATGAATCCATCTGGATGGAAGAC
rs2836493	ACGTTGGATGTTAAGAGTCCGATGCTTGC	ACGTTGGATGGTAATCTGGACTTCTCTTCC
rs2836494	ACGTTGGATGGTGCATTCAATTTGAATTGCTG	ACGTTGGATGCAGTCTTACTTAAACTGAC
rs2836495	ACGTTGGATGGAATTTAACGAACTTCAGC	ACGTTGGATGGGATATTTTCAGGATATCTG
rs2898354	ACGTTGGATGTGTAACAAACCTGCACATCC	ACGTTGGATGGGTACTTTCCAAATATCTGC
rs3065390	ACGTTGGATGCGAGACTCCATCTCAAAAAG	ACGTTGGATGTGGAAAGTACCAATAGCTTC
rs2836496	ACGTTGGATGTGGAGCTTAATGTGTTCTCTG	ACGTTGGATGGTTAGCCATGCATAAGACAG
rs2836497	ACGTTGGATGAGCCGGGATGACTGCTAGAC	ACGTTGGATGAGATGAGGCTGAAGAAGTAA
rs2836498	ACGTTGGATGGGTCTGGGAAAATAGGATG	ACGTTGGATGCACCCTTGCTCTTTCTGAAG
rs2836499	ACGTTGGATGACTAGTCAGAGCACAGTGAG	ACGTTGGATGGCTCTCTCCTTCTTTGACTC
rs2836500	ACGTTGGATGGCTTCTGTTAGTAAGAGG	ACGTTGGATGATCAACTCAGGGCTCTTCTC
rs2836501	ACGTTGGATGACTCACAAAGGTTGACCTTG	ACGTTGGATGGAGGTCCAGGTTGAAAGAAC
rs2836502	ACGTTGGATGGAGGTCCAGGTTGAAAGAAC	ACGTTGGATGACTCACAAAGGTTGACCTTG
rs2836503	ACGTTGGATGGAGCAATTATCAACCCTACG	ACGTTGGATGATTCTCCCCCTTCACTCTTG
rs2836504	ACGTTGGATGGAGTCTGGGTATGGAAAGAG	ACGTTGGATGTTCTAGAAATGGTGTCTGC
rs3787917	ACGTTGGATGTTTGGAGGAGGAATGCCTTG	ACGTTGGATGCGCCACAAACCTAAGAGAA
rs2836505	ACGTTGGATGTTTCTGACTGCTCCACTCTG	ACGTTGGATGGCTCTCCCTCATTGTTCTTC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2836506	ACGTTGGATGGGCTAAGGGCATCATTTTATC	ACGTTGGATGTTTTGCTGATTCATGGATGC
rs2836507	ACGTTGGATGAGCAAAGGTTCTGGTGTGG	ACGTTGGATGAAATGATGCCCTTAGCCCAG
rs2836508	ACGTTGGATGGTGTGATGATATTTTTCTCC	ACGTTGGATGTTTCAGGTATTCCTCTTTGC
rs2836509	ACGTTGGATGTAAGCTTTCTAAGTCAATG	ACGTTGGATGTCATATGATAATGGTCTCTG
rs2836510	ACGTTGGATGCAGGGAGAGATCTAAACAGC	ACGTTGGATGGCCAAAGCTATAACACGTGG
rs2836511	ACGTTGGATGAGAACCTGACTTTTGGAGTG	ACGTTGGATGCTTCCTCATTGGTCAGAGTC
rs2212601	ACGTTGGATGCCAGCCTTTAGAACTGTGAG	ACGTTGGATGTGGGCTGCTGTAACAAAGTG
rs2212602	ACGTTGGATGACTACAACCAGCCAGAGATG	ACGTTGGATGCACAAACCTTGTTGAACCC
rs2226682	ACGTTGGATGCCAAGATTGAACCAGGAAAG	ACGTTGGATGCACAAAAGAATTCAGGAGGTG
rs2836512	ACGTTGGATGCCCCAAAACCTAGCATCCTG	ACGTTGGATGTGTTCTCCCTGCACCTCAAC
rs2836513	ACGTTGGATGCACTGGGGTTAGCAAGAAAC	ACGTTGGATGGACTGTGATTCACCCTGTCT
rs1999328	ACGTTGGATGAGTTACAGCGCAAATTGAGG	ACGTTGGATGGCCTTTATGACTCCATTTCTC
rs2212603	ACGTTGGATGTGGAGGGTGTCTGTGAGTAC	ACGTTGGATGTCATGGAGCAAGGTCTGTGG
rs3787919	ACGTTGGATGCCATCAGCTAGGATTCATGG	ACGTTGGATGTCTGTGAGTACCCACAATG
rs2836514	ACGTTGGATGCAGGTCTAACTAACTGATGAC	ACGTTGGATGGCCTCTACTGTTATTTAAGG
rs1023153	ACGTTGGATGTACAAAAGTGACCTAGAGCC	ACGTTGGATGTTCTTGCAGGACATTGTGCC
rs1023372	ACGTTGGATGCAAATTCAAAAATTCTGTTG	ACGTTGGATGCTCAGAAGTAACATGTACTC
rs2212604	ACGTTGGATGCAGACTTGAGCATATACCAC	ACGTTGGATGACCCATGTGGGAAAATGTTG
rs2226684	ACGTTGGATGGGTGTTGGAAAAGGAACATC	ACGTTGGATGTTAATGATAGTTCCCTCAG
rs2212605	ACGTTGGATGATATGAGTGATTGTCATGGG	ACGTTGGATGTGCATATAAGCTGTCTGCAC
rs2187307	ACGTTGGATGCACATCCTGCAGCTTTAACC	ACGTTGGATGCCTGGCAGTTTCAAGTAACG
rs3065412	ACGTTGGATGGGCTGAGATAGAATGTGCTC	ACGTTGGATGTCTCCTGCTTTGTTCTGGAG
rs2898355	ACGTTGGATGGGCTGAGATAGAATGTGCTC	ACGTTGGATGTCTCCTGCTTTGTTCTGGAG
rs2836518	ACGTTGGATGCACTTGTTGCTTCTCCACC	ACGTTGGATGATGCCAACCTTGCTGATGTC
rs3838110	ACGTTGGATGGAAGTAGTGAAGTGTTCCCC	ACGTTGGATGAGCCTCACTGAATCTTAACG
rs2836519	ACGTTGGATGTGTTTCTCCTTCTCACTGGG	ACGTTGGATGAAAGGCTACAGGAACGAGC
rs3827207	ACGTTGGATGTGTAGTCTGCACCTTCACCT	ACGTTGGATGAGCGGCTGCTGAACATAGAT
rs2836520	ACGTTGGATGCCTGCAAAGGTGTTTGCTTC	ACGTTGGATGGCCACCTAATTTTCTCTC
rs2836521	ACGTTGGATGAAGAATAAGAAGCAAACACC	ACGTTGGATGGTTTTAGGGGAAAGGCATAAG
rs2836522	ACGTTGGATGTGCATCTTTGGTTGTGACAG	ACGTTGGATGGCACATCTACTCTTAGCATG
rs2836523	ACGTTGGATGTCTCTCTTTCTTTCCCTAC	ACGTTGGATGACTCTCAGTTATGATTTCTC
rs2836524	ACGTTGGATGGTGTGTTGGTAGAAACGTTT	ACGTTGGATGGTCACCCCTTCAGATAATAAG
rs2836525	ACGTTGGATGCAGAGCCGAAAACATAGTTC	ACGTTGGATGGTGTGTTGGTAGAAACGTTT
rs3833350	ACGTTGGATGGTTGTTCTTTTGTCTTCTAG	ACGTTGGATGGAATCATGTCCTTCAGTAAGC
rs2836526	ACGTTGGATGATTGTGTCCTGTCCTGCTAG	ACGTTGGATGGACGGCTAGAAGACAAAAGG
rs2836527	ACGTTGGATGGTGTGTTTATGTTCTAGCAGG	ACGTTGGATGGATGCCTTTAGGCAAACATG
rs3834676	ACGTTGGATGAAGCTGAAAAGGATGTGCAG	ACGTTGGATGACAGGGCATACTTCTCTATC
rs2836528	ACGTTGGATGCCAAAACCTCATGCGATCTGC	ACGTTGGATGTGGCGCTGAAGTACTCAATG
rs3761364	ACGTTGGATGAAACAGCACAGCTACCATT	ACGTTGGATGATGAGAAAATGTGTGTGGAG
rs2836529	ACGTTGGATGAGCGGTGTTTTAAATGTCC	ACGTTGGATGCAGAGCCCCAAAAAATTTGG
rs2836530	ACGTTGGATGACAGACAGTGGTCAGAACAT	ACGTTGGATGAAAGATGCCTATAATCCAGG
rs3761366	ACGTTGGATGCAGGTGATAAAAAGCAAGTG	ACGTTGGATGGCCATCAGTTCTTTTTTGGC
rs2836531	ACGTTGGATGGCCTTCGAAAATGTCTCAAG	ACGTTGGATGCACTTGCTTTTTATCACCTG
rs2836532	ACGTTGGATGGAAGACAGCCTTCGAAAATG	ACGTTGGATGCAATGGCTCTTTCAGTAAC
rs2836533	ACGTTGGATGTTTCTGACCTCTCACGGTAC	ACGTTGGATGTGCAGATCTGGAGGTAGATG
rs2836534	ACGTTGGATGAGAAGAGGCTGGGAGAGGAT	ACGTTGGATGTGCTGCTCTTAGGATAAGGG
rs2836535	ACGTTGGATGACAGGAGGAGTTGAGTGTTG	ACGTTGGATGTAGAGGCACGGAGAAGATAG

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2836536	ACGTTGGATGAAAAGCATGGGTACAGGAGG	ACGTTGGATGTAGAGGCACGGAGAAGATAG
rs3827208	ACGTTGGATGGAGGATGAGAGGTACCTGAG	ACGTTGGATGGGGATGATCAAACGTAGT
rs715860	ACGTTGGATGTTCTGGTGGAGGTTTCTTGG	ACGTTGGATGCGAGACATGATCTCAAACCC
rs717231	ACGTTGGATGCAAGAGACTCAAACAGTTGC	ACGTTGGATGTCATAGAAGTTACAGCAGCC
rs2836537	ACGTTGGATGTTGGTGTGTGATCACTCTGG	ACGTTGGATGGAACCTAAGTTTCTCCCAGC
rs2836538	ACGTTGGATGGGTTAGAGCTTACGTAATTC	ACGTTGGATGCTACTTGTGTCACTTCTTTG
rs2836539	ACGTTGGATGTTATCCTCCAAGAGCCTTAG	ACGTTGGATGGGGCAAATGGAGTTCTTATT
rs2836540	ACGTTGGATGCCCAGTTGGTATCAGTGTTG	ACGTTGGATGTGCTGAACATCGTTTGGAGG
rs2836541	ACGTTGGATGCTTGCCTGACACCTTTGTG	ACGTTGGATGGTACTGGCGAAGACATGATG
rs2836542	ACGTTGGATGAGATGAGCCATTTCTACTG	ACGTTGGATGCAGCATGAGAACTGAATGC
rs2836543	ACGTTGGATGAAATGGACTTCTTCAGTAGG	ACGTTGGATGGATACAATTCAACCCATAGC
rs881837	ACGTTGGATGAATGGATGTGGCTCTTGAGG	ACGTTGGATGTATGGAGGGACTTACGAAAG
rs3949052	ACGTTGGATGTTTTCAACGAAACAGATGC	ACGTTGGATGCCAAGTAAATATTCAATCCCC
rs2065307	ACGTTGGATGTTTTCAACGAAACAGATGC	ACGTTGGATGCCAAGTAAATATTCAATCCCC
rs3216105	ACGTTGGATGACCACCATGCCTGGCTAATT	ACGTTGGATGGGCCTGGACAAAATAGTGAG
rs2073427	ACGTTGGATGTTTTGCTTGGGTGTTCTGCC	ACGTTGGATGGGATTTACACTGGTGTTGGG

TABLE 30

dbSNP rs#	Extend Primer	Term Mix
Rs2898353	TCCTGTCTTCAGTGCTTGATTCTG	CGT
rs960818	AGTAGATAACATAAAAGTAACCAGC	ACT
rs960819	GCTATTACCCCTAGCTGTACATAG	ACT
Rs2410034	AAATGTAGCTGTAGTATCTTGAA	ACT
Rs2836437	TTCACACTCAACAACAAACACA	ACT
Rs2836438	TGGAAAGTAAGCTAGACCAAACAG	ACT
Rs2836439	GTATAAAGTGATGCTGCTTGC	ACT
Rs2836440	AACAATTGGGATATGTCTCTCCAC	ACG
Rs2226683	GAGAGTTAATGTGCCCTACTT	ACT
Rs2836441	TAATAGTGCTGGCCATAATGC	ACT
Rs2836442	CTCTAGGCTTACAGTAAACAC	ACT
Rs2836443	TATAAGTTCAGGGTCACAGGTC	ACT
Rs2836444	TGTGTTCTTGGGGTCGCCT	ACT
Rs3787906	TAATGTAGGTGCTGAGAACTTAG	ACT
Rs3838108	GGCTGATTAAAATTCTGTTTCCCC	ACT
Rs2836445	AGACGCAGTAAACTTATGGAT	ACG
Rs2836446	GCCTTGTCCTATCAAGAGCCAAAG	CGT
Rs3787908	CATACAGTAGCTGTGGACAGC	ACT
Rs2836447	ATGTATTACATTGAGAACCATGTG	ACT
Rs2836448	TGTATAGGGCAGGGATAAAGAC	ACT
Rs2836450	AACAACAAATTTACTGATATCATC	ACT
Rs2836451	CTGTCACCCATTGACCTCAC	ACT
Rs1015022	CTTTTATCTGCAGTTGCACCC	ACT
Rs2836452	CGGGAAGATGGCTGCCTTC	ACG

dbSNP rs#	Extend Primer	Term Mix
Rs2836453	CCAAGATGGCCCAGTAGGA	CGT
Rs3787909	AAATAGTAAAATAAAAAAGAGCTCC	ACG
Rs2836454	CACAACCTCCCAAATGAATAAATC	ACT
Rs2836455	TGCTGTGACATTTTAGTGCTTCTG	CGT
Rs2155718	CTCACACACAGCTGGAGTTTA	ACT
Rs2836456	CGTTCTGAAGGTTTTGTGTACA	ACT
Rs2836457	GAGTCACCCGTCCCCTAGA	ACT
Rs2836458	ACAGAAGAGCCAGCCGACA	ACT
Rs2032323	TGCACACTCACTGAAGCCC	ACT
Rs2051400	AAACACTATGTGACGCCACC	ACT
Rs2836459	AGAATGTTACTTTCTGGATTCTAC	ACT
Rs2836460	ATTGTAATTCTCCGTAAAACCC	ACG
Rs2836461	TACCCACACGGAATTCTCATCTAC	ACT
Rs2836462	TCCGTATGTCTCCATCCATCTCA	ACT
Rs2836463	AAACTTAAATTGCTTTAATCAGCT	ACT
Rs2836464	AATATCTTATCACTGCTCCTGTCT	ACG
Rs2836465	GCCCACTTTTGTGTTTGCTTAG	ACT
Rs2836466	TTTGCCCACTTTTGTGTTTGCT	ACT
Rs2836467	TTAATTTTCTTGTCTCTTTCTGTA	ACT
Rs3827204	CCCTCACATCTTCCCCGC	ACT
Rs2836468	GCAGGAAAGAGCATGGGCATTAAC	ACT
Rs3787911	TACATCCAAAAGCCTGCCAG	ACT
Rs2836469	TCCTGCGAGATCCTGCTCA	ACG
Rs2836470	ACAAGCTTAATGTTTTGTTGAGA	ACT
Rs2212599	TTCCCCAACAATAGTCAGAAAA	ACT
Rs2836472	TTCTCTATCATGATGCAGTCC	ACT
Rs2836473	GATGATGAACAGGGCTGTGA	ACG
Rs1888469	AAGGGTCTGAAGAGGAGGC	ACT
Rs1888470	GTTTTCTGCCTCTGATCCTCA	ACT
Rs2032322	CCTATAGGTAACGTGGCTTCT	ACT
Rs2410035	AGGCAGAAGTTGCAGTGAAC	ACG
Rs1573332	GAGAGGCCAGAAAGCCTTC	CGT
Rs2836474	GCACAGGAGAGTCCTCAATT	ACG
Rs2836475	CATGGGAAGCTGCTGAACTA	ACT
Rs3787914	ACAGTGTTTGAGCCCTCCTT	ACT
Rs1888471	AACTGACAGAAGAAAGAAAAATAT	ACG
Rs1888472	TGTGTTGGTGTATAAATCAAGATT	ACG
Rs1888473	CAGTTCTCAGCCAGACGATC	ACG
Rs1888474	GAGTCCAGGGTGCTAATTTT	ACG
Rs2836476	GGTGTTAGCCCTGGGTTCTAATAA	ACG
Rs3787916	TCTCTTATGTAAATACAAAGACG	CGT
Rs2836477	CCTCTTAAATAGCCTGCCTTCA	ACT
rs970043	GCTCCTTGACTCAAGTATTTT	ACG
Rs2212600	AAAACAACTTTCTCTCCCAAAC	ACG

dbSNP rs#	Extend Primer	Term Mix
Rs2836478	CTTGCTTATCTTCAAGCAGTC	CGT
Rs2836479	CCTAATAGGTGTGAAGTGTA AAA	CGT
Rs1475877	CTCCCCGTGTTCTGCATGC	ACG
Rs2836480	CCCACTGTACATCTTACACTC	ACT
Rs2836481	TCCCCCTGAAATCCCATAGC	ACT
Rs2836483	AGGTAATCTCCAACCAAACCT	ACT
Rs2836484	AGTCATCAAGCCATATCTCCA	ACG
Rs2836485	CTCCTCTGGGACGTCAGC	ACT
Rs2836486	CCTCTAAGTTTAGTGGTGGAT	ACT
Rs2836487	TGTTGGGTTCTACACATTCAA	ACT
Rs1893199	CAGACCATGGTTTTGAATGTG	ACG
Rs2836488	GTAGAACCCAACACAGAGCC	ACG
Rs1893200	AGTCTTCGCATAACCAAAACAGA	ACT
Rs1893201	CGCATAACCAAAACAGAAAAGAAC	ACT
Rs2836489	CAAGAGCTCTTTTCAATTCCAG	ACT
Rs1888475	GACATCAAATGATTCCCCTGT	ACT
Rs2836490	GAGCCAAAGCTTTCCTGATG	ACT
Rs2836491	GTGGAAAGGGCACTGTGGT	ACT
Rs2836492	GGCATGGAAGAGCAAGCATC	ACT
Rs2836493	TCCGATGCTTGCTCTTCCAT	ACT
Rs2836494	TGAAGTTTCGTTAAATTCAC TACA	ACT
Rs2836495	CTTCAGCAATTCAAATGAATGCAC	ACT
Rs2898354	TCCGGCACATATATCCTGGAAC	ACT
Rs3065390	AAACAAACAAACAAAAACAGTGTA	ACT
Rs2836496	GTGTTCTGATGTTTCTGGAGT	CGT
Rs2836497	CTGCTAGACATTGTCAGTCC	ACT
Rs2836498	AATAGGATGAGTCAAAGAAGGAG	ACT
Rs2836499	GAGAAGAGCCCTGAGTTGATAAA	ACT
Rs2836500	AGAGGATGAGCAATTCAGGGA	ACT
Rs2836501	CAAAGGTTGACCTTGTTTTCTAT	CGT
Rs2836502	AAGAACTTACATTTTATGGCTTC	ACT
Rs2836503	GATTTGGGAGCAAGGGAGC	ACT
Rs2836504	AGAGTTAAAGATGACTCTAGGCTC	ACT
Rs3787917	GCAGCCAGAGTGGAGCAGT	ACG
Rs2836505	AAGGCATTCTCCTCCAAATCAC	ACT
Rs2836506	GAAATCAAATCAGTTTCTACAAC	ACT
Rs2836507	GTGTTGGAATATTGTTGCCT	ACT
Rs2836508	ATTCTCTACCATTTTATTCTCTTT	ACT
Rs2836509	TTTCTAAGTCAATGTAGGCAAC	ACT
Rs2836510	CAGCTAGTTATCTTACTTCACC	ACT
Rs2836511	AGCAGGTGACAACCCAGACAT	ACT
Rs2212601	TAAGTTTCTGTTGTTTATATGCCA	ACT
Rs2212602	CCAGCCAGAGATGGGATCA	ACG
Rs2226682	GATTGAACCAGGAAAGAAATAGTT	CGT

dbSNP rs#	Extend Primer	Term Mix
Rs2836512	AATGCCAGTTGCCATAGGATA	ACG
Rs2836513	ATAAGAAGATGAGTACTATTATTG	ACT
Rs1999328	ATTGAGGGAAGAGTAAATGATTTC	CGT
Rs2212603	TGTCTGTGAGTACCCCAACAATGAA	ACT
Rs3787919	TCTGTGGCTTCAATGCTGGG	ACT
Rs2836514	ACAGACTTTAACAAAATCACTGA	ACT
Rs1023153	GGGTCATCTCCTTACCTGTCCAA	ACG
Rs1023372	TTCCAAAATTCTGGTTGTGTTTT	ACT
Rs2212604	CTGCCCTTATACATACATAGCTTC	ACG
Rs2226684	AAAAACAATCTGCACAACAAATAT	ACT
Rs2212605	GCAGTGAATATGAACAAAAAAAAA	ACT
Rs2187307	CAGCTTTAACCTCACTCCAC	ACT
Rs3065412	AGTTACAAATCAGGTGGTGCTGG	ACT
Rs2898355	GTTACAAATCAGGTGGTGCTG	ACT
Rs2836518	TAGGAATCGGAGTCAATAATTTT	ACT
Rs3838110	GCTGCACAATCCCCCCCC	CGT
Rs2836519	CCTTCTCACTGGGTTCTCTG	ACG
Rs3827207	TATCACCCCTGTGTCTCTGC	ACG
Rs2836520	CACAAATAGATTATATATCCTGTT	ACT
Rs2836521	AATAAGAAGCAAACACCTTTGCA	ACT
Rs2836522	CCACCCCTTCAGAGAGTTG	ACT
Rs2836523	TCATATTGGTTGATCGTATTGGTT	ACT
Rs2836524	GATTTCAAGGAATGAACTATGTTTT	ACG
Rs2836525	AGCCGAAAACATAGTTCATTCTCTG	ACT
Rs3833350	CTTTTGTCTTCTAGCCGTCAG	ACT
Rs2836526	AGAACATAAAACACAGAAATGCA	ACT
Rs2836527	TTATGTTCTAGCAGGACAGGA	CGT
Rs3834676	AAAAGGATGTGCAGATCGCAT	ACT
Rs2836528	ATCTGCACATCCTTTTCAGCTT	ACG
Rs3761364	CTACCATTCAATTGAGTACTTCAG	ACG
Rs2836529	CTTCAAAATGTGGGTTGATACC	ACT
Rs2836530	GGTCAGAACATGCTGCTTTAT	ACT
Rs3761366	GTGATGGCTTCTAAAAATGTAAA	ACG
Rs2836531	GCATTTGTTACTGCAAAGAGCCAT	ACG
Rs2836532	AGCCTTCGAAAATGTCTCAAG	CGT
Rs2836533	CACACCCATTCCAACCCAAT	ACG
Rs2836534	GCTGAAGGTTTCTGGGAGCA	ACG
Rs2836535	GAGGAGTTGAGTGTGGAACCA	ACG
Rs2836536	ATGGGTACAGGAGGAGTTGA	ACT
Rs3827208	CACCCACCCCAATCACCC	ACT
rs715860	CTTGGTTATCCTTCAGTTTCCA	ACT
rs717231	CTCATTTAGTTTATGTCTTGTTG	ACT
Rs2836537	GCTCATACGCCCTTGGTCTCTAAT	ACT
Rs2836538	AGCTTACGTAATTCAAATCAAGT	ACT

dbSNP rs#	Extend Primer	Term Mix
Rs2836539	TTACACATTTGCACAATGAGGATA	CGT
Rs2836540	GTATCAGTGTGAATGACTGGT	ACT
Rs2836541	TGACACCTTTGTGAATTGCTGAAC	ACT
Rs2836542	CCATTTCCTACTGAAGAAGTCCA	ACT
Rs2836543	CTTCTTCAGTAGGAAATGGCT	ACG
rs881837	GGCTCTTGAGGCCATGCC	ACG
Rs3949052	ACAATTTCTCATGTTGTAAGGATT	ACG
Rs2065307	GGAAACAGATGCCATTTACAATTT	ACG
Rs3216105	GCCTGGCTAATTTTTAAAAAAAAA	CGT
Rs2073427	CTGCCCCACATGACCCA	ACG

Genetic Analysis

[0254] Allelotyping results from the discovery cohort are shown for cases and controls in Table 31. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where “AF” is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs2898353 has the following case and control allele frequencies: case A1 (A) = 0.79; case A2 (T) = 0.21; control A1 (A) = 0.81; and control A2 (T) = 0.19, where the nucleotide is provided in paranthesis. Some SNPs are labeled “untyped” because of failed assays.

TABLE 31

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2898353	231	38783681	A/T	0.21	0.19	0.560
rs960818	882	38784332	A/G	0.59	0.57	0.330
rs960819	960	38784410	A/C	0.13	0.09	0.101
rs2410034	1194	38784644	A/C			
rs2836437	1530	38784980	A/G	0.14	0.14	0.956
rs2836438	1673	38785123	A/G	0.79	0.75	0.077
rs2836439	2096	38785546	C/T	0.70	0.71	0.508
rs2836440	2285	38785735	A/G	0.19	0.18	0.623
rs2226683	5873	38789323	C/T	0.79	0.76	0.312
rs2836441	7256	38790706	A/G	0.12	0.12	0.765
rs2836442	7988	38791438	A/G	0.31	0.30	0.746
rs2836443	8222	38791672	G/T	0.22	0.23	0.728
rs2836444	8381	38791831	C/T	0.19	0.20	0.807
rs3787906	8814	38792264	C/T	0.97	untyped	NA
rs3838108	8915	38792365	-/C	0.58	0.56	0.425
rs2836445	9642	38793092	A/G	0.32	0.35	0.190
rs2836446	9902	38793352	A/T	0.12	0.14	0.274
rs3787908	10619	38794069	A/G			
rs2836447	10927	38794377	C/T	0.68	0.67	0.816
rs2836448	11032	38794482	C/T	0.12	0.14	0.235
rs2836450	14377	38797827	C/T	0.70	0.68	0.460
rs2836451	15608	38799058	C/T	0.92	0.95	0.157

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1015022	15928	38799378	C/G	0.31	0.36	0.072
rs2836452	16296	38799746	A/G	0.18	0.18	0.822
rs2836453	17598	38801048	A/T	0.02	0.02	0.836
rs3787909	19272	38802722	A/G	0.06	0.03	0.091
rs2836454	20084	38803534	A/G	0.04	0.03	0.397
rs2836455	20577	38804027	A/T	0.17	0.13	0.050
rs2155718	28051	38811501	A/G	0.78	0.78	0.950
rs2836456	29466	38812916	A/G	0.94	0.92	0.569
rs2836457	29530	38812980	C/T			
rs2836458	29987	38813437	A/G	0.48	0.46	0.455
rs2032323	30012	38813462	C/T			
rs2051400	30322	38813772	G/T	0.03	NA	NA
rs2836459	32216	38815666	C/T	0.19	0.17	0.319
rs2836460	32516	38815966	C/T			
rs2836461	32544	38815994	A/G			
rs2836462	32746	38816196	A/G			
rs2836463	33137	38816587	G/T	0.67	0.72	0.032
rs2836464	33538	38816988	A/G	0.67	0.67	0.991
rs2836465	33798	38817248	C/T			
rs2836466	33802	38817252	A/C	0.39	0.40	0.627
rs2836467	33964	38817414	C/T			
rs3827204	34132	38817582	A/G	0.45	0.42	0.213
rs2836468	34210	38817660	C/T	0.13	0.14	0.678
rs3787911	34317	38817767	A/G	0.13	0.12	0.862
rs2836469	34499	38817949	C/T	0.38	0.40	0.250
rs2836470	34753	38818203	A/C	0.73	0.74	0.939
rs2212599	34845	38818295	C/T	0.66	0.64	0.474
rs2836472	35335	38818785	C/T	0.40	0.35	0.071
rs2836473	36423	38819873	C/T	0.53	0.54	0.755
rs1888469	36450	38819900	A/G	0.45	0.49	0.175
rs1888470	36481	38819931	G/T	0.17	0.18	0.623
rs2032322	38447	38821897	C/G	0.50	0.50	0.879
rs2410035	38784	38822234	C/T			
rs1573332	39387	38822837	A/T	0.57	0.58	0.609
rs2836474	39458	38822908	C/T	0.33	0.35	0.564
rs2836475	39822	38823272	C/G	0.17	0.14	0.113
rs3787914	40305	38823755	C/G	0.73	0.73	0.987
rs1888471	40869	38824319	C/T	0.29	0.26	0.175
rs1888472	40926	38824376	C/T	0.62	0.63	0.818
rs1888473	41010	38824460	C/T	0.63	0.65	0.435
rs1888474	41134	38824584	C/T	0.28	0.23	0.099
rs2836476	41984	38825434	A/G	0.46	0.44	0.379
rs3787916	42172	38825622	A/T	0.45	0.43	0.314
rs2836477	42753	38826203	G/T	0.94	0.96	0.196
rs970043	43011	38826461	C/T	0.04	0.04	0.549
rs2212600	43176	38826626	A/G			
rs2836478	43320	38826770	G/T	0.76	0.75	0.914
rs2836479	43381	38826831	A/T	0.44	0.43	0.670
rs1475877	44142	38827592	A/G	0.35	0.32	0.110
rs2836480	44383	38827833	A/G	0.46	0.43	0.153
rs2836481	44726	38828176	C/T	0.42	0.40	0.434
rs2836483	45087	38828537	A/G	0.47	0.45	0.393
rs2836484	45141	38828591	C/T	0.46	0.47	0.671
rs2836485	45359	38828809	C/G	0.16	0.17	0.643
rs2836486	45421	38828871	C/T			
rs2836487	45456	38828906	C/T	0.02	0.03	0.758
rs1893199	45467	38828917	C/T	0.62	0.65	0.220

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2836488	45486	38828936	C/T	0.25	0.23	0.360
rs1893200	45709	38829159	A/G	0.16	0.14	0.177
rs1893201	45716	38829166	A/G	0.84	0.87	0.060
rs2836489	47626	38831076	C/T	0.29	0.31	0.502
rs1888475	49413	38832863	A/G			
rs2836490	49796	38833246	C/T	0.94	0.93	0.731
rs2836491	49962	38833412	A/G	0.10	0.08	0.219
rs2836492	50075	38833525	C/T	0.20	0.22	0.518
rs2836493	50093	38833543	A/G	0.95	0.94	0.850
rs2836494	50571	38834021	C/T	0.72	0.70	0.536
rs2836495	50615	38834065	A/G	0.82	0.78	0.142
rs2898354	50780	38834230	A/G	0.25	0.25	0.728
rs3065390	50851	38834301	-T/A	0.10	0.11	0.845
rs2836496	51459	38834909	A/C	0.80	0.84	0.064
rs2836497	53193	38836643	C/T	0.65	0.65	0.935
rs2836498	53702	38837152	C/T	0.43	0.44	0.682
rs2836499	53736	38837186	A/C	0.33	0.30	0.169
rs2836500	53795	38837245	C/T			
rs2836501	54109	38837559	A/T	0.36	0.34	0.234
rs2836502	54126	38837576	C/T	0.31	0.29	0.427
rs2836503	54230	38837680	A/C	0.32	0.29	0.194
rs2836504	54894	38838344	C/T	0.51	0.54	0.170
rs3787917	55455	38838905	A/G	0.56	0.60	0.137
rs2836505	55499	38838949	A/G	0.73	0.78	0.022
rs2836506	56522	38839972	C/T	0.52	0.56	0.145
rs2836507	56662	38840112	C/T	0.51	0.54	0.173
rs2836508	56954	38840404	A/G	0.53	0.56	0.376
rs2836509	57267	38840717	A/G	0.35	0.31	0.089
rs2836510	58282	38841732	A/G	0.65	0.59	0.034
rs2836511	58916	38842366	A/C	0.32	0.30	0.315
rs2212601	59544	38842994	C/G	0.45	0.46	0.568
rs2212602	59666	38843116	C/T	0.30	0.28	0.644
rs2226682	59913	38843363	A/T	0.38	0.35	0.164
rs2836512	66846	38850296	A/G	0.94	0.94	0.896
rs2836513	67245	38850695	G/T	0.23	0.22	0.713
rs1999328	67652	38851102	A/C	0.79	0.79	0.973
rs2212603	67955	38851405	A/G	0.73	0.72	0.776
rs3787919	67966	38851416	A/C			
rs2836514	68420	38851870	A/G	0.52	0.54	0.319
rs1023153	70226	38853676	A/G	0.09	0.09	0.985
rs1023372	70810	38854260	C/T	0.83	0.81	0.518
rs2212604	72246	38855696	A/G	0.68	0.71	0.237
rs2226684	73330	38856780	G/T	0.83	0.81	0.462
rs2212605	73457	38856907	C/T	0.82	0.85	0.255
rs2187307	74389	38857839	A/G	0.13	0.13	0.869
rs3065412	74638	38858088	-T/A			
rs2898355	74640	38858090	A/C	0.96	0.94	0.413
rs2836518	75358	38858808	A/C	0.10	0.12	0.261
rs3838110	75952	38859402	-G	0.66	0.67	0.790
rs2836519	76098	38859548	A/G	0.60	0.61	0.509
rs3827207	77836	38861286	A/G	0.62	0.63	0.575
rs2836520	78449	38861899	A/C			
rs2836521	78507	38861957	G/T	0.07	0.08	0.551
rs2836522	80031	38863481	G/T	0.11	0.08	0.155
rs2836523	81695	38865145	C/T			
rs2836524	82775	38866225	A/G	0.05	0.04	0.321
rs2836525	82795	38866245	A/G	0.11	0.11	0.875

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs3833350	84611	38868061	-/C			
rs2836526	84657	38868107	C/T	0.83	0.86	0.292
rs2836527	84693	38868143	A/C	0.08	0.08	0.936
rs3834676	85020	38868470	-/T	0.80	0.83	0.191
rs2836528	85048	38868498	C/T	0.84	0.87	0.089
rs3761364	85100	38868550	C/T	0.06	0.04	0.159
rs2836529	85325	38868775	A/C	0.09	0.06	0.100
rs2836530	85452	38868902	C/T			
rs3761366	85868	38869318	A/G	0.06	0.04	0.179
rs2836531	85936	38869386	A/G	0.49	0.50	0.729
rs2836532	85990	38869440	A/T	0.30	0.29	0.766
rs2836533	86139	38869589	C/T	0.47	0.48	0.751
rs2836534	86497	38869947	C/T	0.87	0.87	0.874
rs2836535	87236	38870686	A/G	0.93	0.92	0.628
rs2836536	87248	38870698	C/T	0.86	0.84	0.474
rs3827208	87533	38870983	C/G	0.51	0.53	0.459
rs715860	87912	38871362	A/G	0.08	0.09	0.627
rs717231	88108	38871558	G/T	0.65	0.67	0.382
rs2836537	88494	38871944	A/C	0.43	0.40	0.239
rs2836538	89598	38873048	A/C			
rs2836539	90235	38873685	A/T	0.98	0.97	0.796
rs2836540	91287	38874737	G/T			
rs2836541	91359	38874809	C/T	0.07	0.06	0.403
rs2836542	92384	38875834	A/C	0.36	0.38	0.418
rs2836543	92410	38875860	C/T	0.54	0.50	0.202
rs881837	92900	38876350	C/T	0.29	0.28	0.639
rs3949052	94495	38877945	A/G			
rs2065307	94512	38877962	A/G			
rs3216105	97777	38881227	-/A	0.32	0.28	0.265
rs2073427	98333	38881783	C/T	0.09	0.07	0.242

[0255] The *ERG* proximal SNPs were also allelotyped in the replication cohorts using the methods described herein and the primers provided in Tables 29 and 30. The replication allelotyping results for replication cohort #1 and replication cohort #2 are provided in Tables 32 and 33, respectively.

TABLE 32

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2898353	231	38783681	A/T	0.19	0.19	0.773
rs960818	882	38784332	A/G	0.59	0.57	0.600
rs960819	960	38784410	A/C	0.07	NA	0.132
rs2410034	1194	38784644	A/C			
rs2836437	1530	38784980	A/G	0.14	0.14	0.957
rs2836438	1673	38785123	A/G	0.80	0.77	0.402
rs2836439	2096	38785546	C/T	0.68	0.73	0.089
rs2836440	2285	38785735	A/G	0.20	0.18	0.421
rs2226683	5873	38789323	C/T	0.78	0.76	0.622
rs2836441	7256	38790706	A/G	0.12	0.12	0.946
rs2836442	7988	38791438	A/G	0.30	0.32	0.674
rs2836443	8222	38791672	G/T	0.22	0.25	0.332
rs2836444	8381	38791831	C/T	0.20	0.20	0.908
rs3787906	8814	38792264	C/T	0.97	untyped	NA

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs3838108	8915	38792365	-/C	0.58	0.56	0.604
rs2836445	9642	38793092	A/G	0.33	0.37	0.211
rs2836446	9902	38793352	A/T	0.13	0.15	0.481
rs3787908	10619	38794069	A/G			
rs2836447	10927	38794377	C/T	0.67	0.67	0.843
rs2836448	11032	38794482	C/T	0.13	0.15	0.521
rs2836450	14377	38797827	C/T	0.67	0.67	0.989
rs2836451	15608	38799058	C/T	0.92	0.95	0.214
rs1015022	15928	38799378	C/G	0.30	0.36	0.076
rs2836452	16296	38799746	A/G	0.18	0.18	0.982
rs2836453	17598	38801048	A/T	0.02	untyped	NA
rs3787909	19272	38802722	A/G	0.06	0.03	0.110
rs2836454	20084	38803534	A/G	0.03	0.03	0.746
rs2836455	20577	38804027	A/T	0.17	0.12	0.080
rs2155718	28051	38811501	A/G	0.78	0.79	0.747
rs2836456	29466	38812916	A/G	0.91	0.91	0.915
rs2836457	29530	38812980	C/T			
rs2836458	29987	38813437	A/G	0.48	0.47	0.626
rs2032323	30012	38813462	C/T			
rs2051400	30322	38813772	G/T	0.02	untyped	NA
rs2836459	32216	38815666	C/T	0.20	0.16	0.278
rs2836460	32516	38815966	C/T			
rs2836461	32544	38815994	A/G			
rs2836462	32746	38816196	A/G			
rs2836463	33137	38816587	G/T	0.67	0.75	0.011
rs2836464	33538	38816988	A/G	0.66	0.68	0.586
rs2836465	33798	38817248	C/T			
rs2836466	33802	38817252	A/C	0.39	0.41	0.507
rs2836467	33964	38817414	C/T			
rs3827204	34132	38817582	A/G	0.45	0.41	0.229
rs2836468	34210	38817660	C/T	0.13	0.14	0.736
rs3787911	34317	38817767	A/G	0.14	0.13	0.856
rs2836469	34499	38817949	C/T	0.37	0.41	0.168
rs2836470	34753	38818203	A/C	0.72	0.73	0.854
rs2212599	34845	38818295	C/T	0.63	0.65	0.636
rs2836472	35335	38818785	C/T	0.41	0.35	0.145
rs2836473	36423	38819873	C/T	0.51	0.54	0.291
rs1888469	36450	38819900	A/G	0.45	0.49	0.281
rs1888470	36481	38819931	G/T	0.17	0.17	0.949
rs2032322	38447	38821897	C/G	0.51	0.53	0.476
rs2410035	38784	38822234	C/T			
rs1573332	39387	38822837	A/T	0.56	0.60	0.279
rs2836474	39458	38822908	C/T	0.33	0.36	0.330
rs2836475	39822	38823272	C/G	0.18	0.13	0.049
rs3787914	40305	38823755	C/G	0.73	0.74	0.977
rs1888471	40869	38824319	C/T	0.31	0.26	0.134
rs1888472	40926	38824376	C/T	0.62	0.65	0.247
rs1888473	41010	38824460	C/T	0.63	0.67	0.210
rs1888474	41134	38824584	C/T	0.28	0.21	0.091
rs2836476	41984	38825434	A/G	0.47	0.44	0.346
rs3787916	42172	38825622	A/T	0.46	0.41	0.171
rs2836477	42753	38826203	G/T	0.94	0.97	0.294
rs970043	43011	38826461	C/T	0.05	0.03	0.331
rs2212600	43176	38826626	A/G			
rs2836478	43320	38826770	G/T	0.75	0.75	0.983
rs2836479	43381	38826831	A/T	0.44	0.43	0.752
rs1475877	44142	38827592	A/G	0.35	0.31	0.166

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2836480	44383	38827833	A/G	0.45	0.41	0.254
rs2836481	44726	38828176	C/T	0.42	0.39	0.330
rs2836483	45087	38828537	A/G	0.46	0.46	0.797
rs2836484	45141	38828591	C/T	0.45	0.47	0.553
rs2836485	45359	38828809	C/G	0.18	0.18	0.993
rs2836486	45421	38828871	C/T			
rs2836487	45456	38828906	C/T	0.03	0.03	0.955
rs1893199	45467	38828917	C/T	0.61	0.67	0.071
rs2836488	45486	38828936	C/T	0.27	0.23	0.246
rs1893200	45709	38829159	A/G	0.16	0.13	0.203
rs1893201	45716	38829166	A/G	0.83	0.89	0.021
rs2836489	47626	38831076	C/T	0.30	0.31	0.702
rs1888475	49413	38832863	A/G			
rs2836490	49796	38833246	C/T	0.94	0.95	0.662
rs2836491	49962	38833412	A/G	0.10	0.06	0.038
rs2836492	50075	38833525	C/T	0.20	0.22	0.651
rs2836493	50093	38833543	A/G	0.93	0.95	0.397
rs2836494	50571	38834021	C/T	0.73	0.71	0.592
rs2836495	50615	38834065	A/G	0.81	0.77	0.212
rs2898354	50780	38834230	A/G	0.24	0.24	0.827
rs3065390	50851	38834301	-T/A	0.10	0.11	0.743
rs2836496	51459	38834909	A/C	0.78	0.86	0.022
rs2836497	53193	38836643	C/T	0.65	0.66	0.733
rs2836498	53702	38837152	C/T	0.44	0.46	0.576
rs2836499	53736	38837186	A/C	0.33	0.29	0.200
rs2836500	53795	38837245	C/T			
rs2836501	54109	38837559	A/T	0.36	0.32	0.167
rs2836502	54126	38837576	C/T	0.31	0.27	0.206
rs2836503	54230	38837680	A/C	0.32	0.28	0.173
rs2836504	54894	38838344	C/T	0.50	0.57	0.033
rs3787917	55455	38838905	A/G	0.56	0.62	0.033
rs2836505	55499	38838949	A/G	0.72	0.81	0.004
rs2836506	56522	38839972	C/T	0.52	0.58	0.093
rs2836507	56662	38840112	C/T	0.51	0.56	0.134
rs2836508	56954	38840404	A/G	0.53	0.58	0.170
rs2836509	57267	38840717	A/G	0.35	0.30	0.136
rs2836510	58282	38841732	A/G	0.62	0.56	0.035
rs2836511	58916	38842366	A/C	0.33	0.30	0.273
rs2212601	59544	38842994	C/G	0.44	0.46	0.675
rs2212602	59666	38843116	C/T	0.29	0.27	0.571
rs2226682	59913	38843363	A/T	0.38	0.33	0.127
rs2836512	66846	38850296	A/G	0.93	0.96	0.261
rs2836513	67245	38850695	G/T	0.23	0.22	0.692
rs1999328	67652	38851102	A/C	0.79	0.80	0.618
rs2212603	67955	38851405	A/G	0.73	0.74	0.676
rs3787919	67966	38851416	A/C			
rs2836514	68420	38851870	A/G	0.51	0.57	0.044
rs1023153	70226	38853676	A/G	0.09	0.09	0.699
rs1023372	70810	38854260	C/T	0.82	untyped	NA
rs2212604	72246	38855696	A/G	0.67	0.73	0.063
rs2226684	73330	38856780	G/T	0.82	0.82	0.992
rs2212605	73457	38856907	C/T	0.83	0.86	0.180
rs2187307	74389	38857839	A/G	0.14	0.13	0.901
rs3065412	74638	38858088	-T/A			
rs2898355	74640	38858090	A/C	0.95	0.93	0.442
rs2836518	75358	38858808	A/C	0.11	0.14	0.248
rs3838110	75952	38859402	-G	0.65	0.68	0.399

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2836519	76098	38859548	A/G	0.59	0.64	0.134
rs3827207	77836	38861286	A/G	0.60	0.64	0.205
rs2836520	78449	38861899	A/C			
rs2836521	78507	38861957	G/T	0.08	0.09	0.765
rs2836522	80031	38863481	G/T	0.12	0.07	0.033
rs2836523	81695	38865145	C/T			
rs2836524	82775	38866225	A/G	0.05	0.04	0.539
rs2836525	82795	38866245	A/G	0.12	0.09	0.179
rs3833350	84611	38868061	-/C			
rs2836526	84657	38868107	C/T	0.83	0.85	0.536
rs2836527	84693	38868143	A/C	0.08	0.07	0.444
rs3834676	85020	38868470	-/T	0.79	0.82	0.270
rs2836528	85048	38868498	C/T	0.82	0.86	0.130
rs3761364	85100	38868550	C/T	0.08	0.05	0.132
rs2836529	85325	38868775	A/C	0.09	0.07	0.214
rs2836530	85452	38868902	C/T			
rs3761366	85868	38869318	A/G	0.07	0.04	0.259
rs2836531	85936	38869386	A/G	0.49	0.50	0.741
rs2836532	85990	38869440	A/T	0.30	0.30	0.921
rs2836533	86139	38869589	C/T	0.48	0.48	0.843
rs2836534	86497	38869947	C/T	0.86	0.89	0.374
rs2836535	87236	38870686	A/G	0.91	0.91	0.933
rs2836536	87248	38870698	C/T	0.86	0.86	0.945
rs3827208	87533	38870983	C/G	0.51	0.55	0.183
rs715860	87912	38871362	A/G	0.07	0.07	0.893
rs717231	88108	38871558	G/T	0.65	0.68	0.506
rs2836537	88494	38871944	A/C	0.43	0.39	0.251
rs2836538	89598	38873048	A/C			
rs2836539	90235	38873685	A/T	0.98	0.98	0.910
rs2836540	91287	38874737	G/T			
rs2836541	91359	38874809	C/T	0.09	0.06	0.324
rs2836542	92384	38875834	A/C	0.37	0.41	0.365
rs2836543	92410	38875860	C/T	0.54	0.55	0.863
rs881837	92900	38876350	C/T	0.30	0.28	0.673
rs3949052	94495	38877945	A/G			
rs2065307	94512	38877962	A/G			
rs3216105	97777	38881227	-/A	0.31	0.29	0.603
rs2073427	98333	38881783	C/T	0.09	0.06	0.249

TABLE 33

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2898353	231	38783681	A/T	0.22	0.21	0.629
rs960818	882	38784332	A/G	0.59	0.55	0.351
rs960819	960	38784410	A/C	0.12	0.01	
rs2410034	1194	38784644	A/C			
rs2836437	1530	38784980	A/G	0.14	0.14	0.989
rs2836438	1673	38785123	A/G	0.78	0.71	0.047
rs2836439	2096	38785546	C/T	0.72	0.68	0.265
rs2836440	2285	38785735	A/G	0.18	0.19	0.789
rs2226683	5873	38789323	C/T	0.80	0.77	0.342
rs2836441	7256	38790706	A/G	0.11	0.12	0.559
rs2836442	7988	38791438	A/G	0.32	0.28	0.269
rs2836443	8222	38791672	G/T	0.23	0.21	0.504

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2836444	8381	38791831	C/T	0.19	0.19	0.829
rs3787906	8814	38792264	C/T	0.97	untyped	
rs3838108	8915	38792365	-/C	0.58	0.55	0.526
rs2836445	9642	38793092	A/G	0.30	0.32	0.722
rs2836446	9902	38793352	A/T	0.11	0.14	0.425
rs3787908	10619	38794069	A/G			
rs2836447	10927	38794377	C/T	0.68	0.68	0.908
rs2836448	11032	38794482	C/T	0.11	0.14	0.302
rs2836450	14377	38797827	C/T	0.73	0.70	0.314
rs2836451	15608	38799058	C/T	0.93	0.94	0.499
rs1015022	15928	38799378	C/G	0.33	0.35	0.527
rs2836452	16296	38799746	A/G	0.17	0.18	0.750
rs2836453	17598	38801048	A/T	0.02	0.02	0.934
rs3787909	19272	38802722	A/G	0.05	0.04	0.546
rs2836454	20084	38803534	A/G	0.05	0.03	0.379
rs2836455	20577	38804027	A/T	0.17	0.15	0.472
rs2155718	28051	38811501	A/G	0.79	0.78	0.704
rs2836456	29466	38812916	A/G	0.97	0.94	0.174
rs2836457	29530	38812980	C/T			
rs2836458	29987	38813437	A/G	0.48	0.45	0.532
rs2032323	30012	38813462	C/T			
rs2051400	30322	38813772	G/T	0.04	0.02	0.476
rs2836459	32216	38815666	C/T	0.19	0.18	0.921
rs2836460	32516	38815966	C/T			
rs2836461	32544	38815994	A/G			
rs2836462	32746	38816196	A/G			
rs2836463	33137	38816587	G/T	0.68	0.68	0.988
rs2836464	33538	38816988	A/G	0.69	0.66	0.430
rs2836465	33798	38817248	C/T			
rs2836466	33802	38817252	A/C	0.39	0.39	0.948
rs2836467	33964	38817414	C/T			
rs3827204	34132	38817582	A/G	0.45	0.43	0.614
rs2836468	34210	38817660	C/T	0.12	0.12	0.879
rs3787911	34317	38817767	A/G	0.12	0.11	0.901
rs2836469	34499	38817949	C/T	0.38	0.39	0.914
rs2836470	34753	38818203	A/C	0.75	0.74	0.960
rs2212599	34845	38818295	C/T	0.71	0.64	0.095
rs2836472	35335	38818785	C/T	0.40	0.36	0.321
rs2836473	36423	38819873	C/T	0.56	0.53	0.433
rs1888469	36450	38819900	A/G	0.45	0.49	0.399
rs1888470	36481	38819931	G/T	0.16	0.19	0.356
rs2032322	38447	38821897	C/G	0.50	0.45	0.190
rs2410035	38784	38822234	C/T			
rs1573332	39387	38822837	A/T	0.58	0.56	0.554
rs2836474	39458	38822908	C/T	0.34	0.33	0.762
rs2836475	39822	38823272	C/G	0.15	0.14	0.817
rs3787914	40305	38823755	C/G	0.73	0.73	0.934
rs1888471	40869	38824319	C/T	0.28	0.27	0.760
rs1888472	40926	38824376	C/T	0.63	0.58	0.302
rs1888473	41010	38824460	C/T	0.63	0.62	0.683
rs1888474	41134	38824584	C/T	0.27	0.26	0.853
rs2836476	41984	38825434	A/G	0.46	0.45	0.838
rs3787916	42172	38825622	A/T	0.44	0.45	0.827
rs2836477	42753	38826203	G/T	0.94	0.95	0.505
rs970043	43011	38826461	C/T	0.04	0.04	0.848
rs2212600	43176	38826626	A/G			
rs2836478	43320	38826770	G/T	0.76	0.75	0.893

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2836479	43381	38826831	A/T	0.44	0.43	0.801
rs1475877	44142	38827592	A/G	0.35	0.33	0.450
rs2836480	44383	38827833	A/G	0.47	0.44	0.444
rs2836481	44726	38828176	C/T	0.41	0.41	0.999
rs2836483	45087	38828537	A/G	0.48	0.44	0.306
rs2836484	45141	38828591	C/T	0.46	0.46	0.939
rs2836485	45359	38828809	C/G	0.15	0.17	0.483
rs2836486	45421	38828871	C/T			
rs2836487	45456	38828906	C/T	NA	0.03	NA
rs1893199	45467	38828917	C/T	0.63	0.62	0.868
rs2836488	45486	38828936	C/T	0.23	0.22	0.913
rs1893200	45709	38829159	A/G	0.17	0.16	0.653
rs1893201	45716	38829166	A/G	0.85	0.85	0.947
rs2836489	47626	38831076	C/T	0.27	0.30	0.597
rs1888475	49413	38832863	A/G			
rs2836490	49796	38833246	C/T	0.94	0.91	0.196
rs2836491	49962	38833412	A/G	0.09	0.11	0.493
rs2836492	50075	38833525	C/T	0.20	0.21	0.669
rs2836493	50093	38833543	A/G	0.96	0.93	0.211
rs2836494	50571	38834021	C/T	0.70	0.69	0.697
rs2836495	50615	38834065	A/G	0.82	0.80	0.510
rs2898354	50780	38834230	A/G	0.27	0.26	0.846
rs3065390	50851	38834301	-/TA	0.11	0.10	0.936
rs2836496	51459	38834909	A/C	0.81	0.80	0.746
rs2836497	53193	38836643	C/T	0.66	0.64	0.756
rs2836498	53702	38837152	C/T	0.41	0.40	0.844
rs2836499	53736	38837186	A/C	0.32	0.30	0.567
rs2836500	53795	38837245	C/T			
rs2836501	54109	38837559	A/T	0.36	0.36	0.917
rs2836502	54126	38837576	C/T	0.31	0.32	0.738
rs2836503	54230	38837680	A/C	0.32	0.31	0.730
rs2836504	54894	38838344	C/T	0.52	0.50	0.620
rs3787917	55455	38838905	A/G	0.57	0.56	0.759
rs2836505	55499	38838949	A/G	0.74	0.74	0.982
rs2836506	56522	38839972	C/T	0.52	0.53	0.907
rs2836507	56662	38840112	C/T	0.51	0.52	0.785
rs2836508	56954	38840404	A/G	0.53	0.52	0.709
rs2836509	57267	38840717	A/G	0.35	0.33	0.453
rs2836510	58282	38841732	A/G	0.68	0.65	0.457
rs2836511	58916	38842366	A/C	0.32	0.31	0.832
rs2212601	59544	38842994	C/G	0.45	0.47	0.717
rs2212602	59666	38843116	C/T	0.30	0.30	0.994
rs2226682	59913	38843363	A/T	0.39	0.38	0.801
rs2836512	66846	38850296	A/G	0.94	0.91	0.184
rs2836513	67245	38850695	G/T	0.23	0.23	0.949
rs1999328	67652	38851102	A/C	0.80	0.77	0.487
rs2212603	67955	38851405	A/G	0.74	0.70	0.289
rs3787919	67966	38851416	A/C			
rs2836514	68420	38851870	A/G	0.53	0.49	0.363
rs1023153	70226	38853676	A/G	0.08	0.09	0.611
rs1023372	70810	38854260	C/T	0.84	0.81	0.315
rs2212604	72246	38855696	A/G	0.69	0.68	0.641
rs2226684	73330	38856780	G/T	0.85	0.81	0.216
rs2212605	73457	38856907	C/T	0.82	0.82	0.927
rs2187307	74389	38857839	A/G	0.12	0.13	0.685
rs3065412	74638	38858088	-/AA			
rs2898355	74640	38858090	A/C	0.96	0.96	0.893

dbSNP rs#	Position in SEQ ID NO: 4	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2836518	75358	38858808	A/C	0.10	0.11	0.823
rs3838110	75952	38859402	-/G	0.68	0.65	0.457
rs2836519	76098	38859548	A/G	0.60	0.57	0.357
rs3827207	77836	38861286	A/G	0.64	0.61	0.449
rs2836520	78449	38861899	A/C			
rs2836521	78507	38861957	G/T	0.06	0.07	0.625
rs2836522	80031	38863481	G/T	0.09	0.10	0.810
rs2836523	81695	38865145	C/T			
rs2836524	82775	38866225	A/G	0.05	0.04	0.419
rs2836525	82795	38866245	A/G	0.10	0.14	0.132
rs3833350	84611	38868061	-/C			
rs2836526	84657	38868107	C/T	0.83	0.86	0.342
rs2836527	84693	38868143	A/C	0.08	0.11	0.209
rs3834676	85020	38868470	-/T	0.81	0.84	0.442
rs2836528	85048	38868498	C/T	0.86	0.88	0.350
rs3761364	85100	38868550	C/T	0.04	0.03	0.643
rs2836529	85325	38868775	A/C	0.08	0.06	0.271
rs2836530	85452	38868902	C/T			
rs3761366	85868	38869318	A/G	0.06	0.04	0.473
rs2836531	85936	38869386	A/G	0.49	0.49	0.915
rs2836532	85990	38869440	A/T	0.31	0.28	0.446
rs2836533	86139	38869589	C/T	0.47	0.48	0.810
rs2836534	86497	38869947	C/T	0.88	0.84	0.149
rs2836535	87236	38870686	A/G	0.94	0.92	0.378
rs2836536	87248	38870698	C/T	0.86	0.82	0.311
rs3827208	87533	38870983	C/G	0.51	0.49	0.598
rs715860	87912	38871362	A/G	0.09	0.11	0.463
rs717231	88108	38871558	G/T	0.65	0.67	0.588
rs2836537	88494	38871944	A/C	0.42	0.41	0.694
rs2836538	89598	38873048	A/C			
rs2836539	90235	38873685	A/T	0.97	0.97	0.749
rs2836540	91287	38874737	G/T			
rs2836541	91359	38874809	C/T	0.05	0.05	0.895
rs2836542	92384	38875834	A/C	0.34	0.34	0.998
rs2836543	92410	38875860	C/T	untyped	0.43	NA
rs881837	92900	38876350	C/T	0.29	0.28	0.811
rs3949052	94495	38877945	A/G			
rs2065307	94512	38877962	A/G			
rs3216105	97777	38881227	-/A	0.32	0.28	0.273
rs2073427	98333	38881783	C/T	0.08	0.07	0.700

[0256] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1D for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1D can be determined by consulting Table 31. For example, the left-most X on the left graph is at position 38783681. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0257] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than 10^{-8} were truncated at that value.

[0258] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

Example 8

In Vitro Production of Target Polypeptides

[0259] cDNA is cloned into a pIVEX 2.3-MCS vector (Roche Biochem) using a directional cloning method. A cDNA insert is prepared using PCR with forward and reverse primers having 5' restriction site tags (in frame) and 5-6 additional nucleotides in addition to 3' gene-specific portions, the latter of which is typically about twenty to about twenty-five base pairs in length. A Sal I restriction site is introduced by the forward primer and a Sma I restriction site is introduced by the reverse primer. The ends of PCR products are cut with the corresponding restriction enzymes (*i.e.*, Sal I and Sma I) and the products are gel-purified. The pIVEX 2.3-MCS vector is linearized using the same restriction enzymes, and the fragment with the correct sized fragment is isolated by gel-purification. Purified PCR product is ligated into the linearized pIVEX 2.3-MCS vector and *E. coli* cells transformed for plasmid amplification. The newly constructed expression vector is verified by restriction mapping and used for protein production.

[0260] *E. coli* lysate is reconstituted with 0.25 ml of Reconstitution Buffer, the Reaction Mix is reconstituted with 0.8 ml of Reconstitution Buffer; the Feeding Mix is reconstituted with 10.5 ml of Reconstitution Buffer; and the Energy Mix is reconstituted with 0.6 ml of Reconstitution Buffer. 0.5 ml of the Energy Mix was added to the Feeding Mix to obtain the Feeding Solution. 0.75 ml of Reaction Mix, 50 µl of Energy Mix, and 10 µg of the template DNA is added to the *E. coli* lysate.

[0261] Using the reaction device (Roche Biochem), 1 ml of the Reaction Solution is loaded into the reaction compartment. The reaction device is turned upside-down and 10 ml of the Feeding Solution is loaded into the feeding compartment. All lids are closed and the reaction device is loaded into the RTS500 instrument. The instrument is run at 30°C for 24 hours with a stir bar speed of 150 rpm. The pIVEX 2.3 MCS vector includes a nucleotide sequence that encodes six consecutive histidine amino acids on the C-terminal end of the target polypeptide for the purpose of protein purification. Target polypeptide is purified by contacting the contents of reaction device with resin modified with Ni²⁺ ions. Target polypeptide is eluted from the resin with a solution containing free Ni²⁺ ions.

Example 9

Cellular Production of Target Polypeptides

[0262] Nucleic acids are cloned into DNA plasmids having phage recombination sites and target polypeptides are expressed therefrom in a variety of host cells. Alpha phage genomic DNA contains short sequences known as attP sites, and *E. coli* genomic DNA contains unique, short sequences known as attB sites. These regions share homology, allowing for integration of phage DNA into *E. coli* via directional, site-specific recombination using the phage protein Int and the *E. coli* protein IHF. Integration produces two new att sites, L and R, which flank the inserted prophage DNA. Phage excision from *E. coli* genomic DNA can also be accomplished using these two proteins with the addition of a second phage protein, Xis. DNA vectors have been produced where the integration/excision process is modified to allow for the directional integration or excision of a target DNA fragment into a backbone vector in a rapid *in vitro* reaction (Gateway™ Technology (Invitrogen, Inc.)).

[0263] A first step is to transfer the nucleic acid insert into a shuttle vector that contains attL sites surrounding the negative selection gene, ccdB (e.g. pENTER vector, Invitrogen, Inc.). This transfer process is accomplished by digesting the nucleic acid from a DNA vector used for sequencing, and to ligate it into the multicloning site of the shuttle vector, which will place it between the two attL sites while removing the negative selection gene ccdB. A second method is to amplify the nucleic acid by the polymerase chain reaction (PCR) with primers containing attB sites. The amplified fragment then is integrated into the shuttle vector using Int and IHF. A third method is to utilize a topoisomerase-mediated process, in which the nucleic acid is amplified via PCR using gene-specific primers with the 5' upstream primer containing an additional CACC sequence (e.g., TOPO® expression kit (Invitrogen, Inc.)). In conjunction with Topoisomerase I, the PCR amplified fragment can be cloned into the shuttle vector via the attL sites in the correct orientation.

[0264] Once the nucleic acid is transferred into the shuttle vector, it can be cloned into an expression vector having attR sites. Several vectors containing attR sites for expression of target polypeptide as a

native polypeptide, N-fusion polypeptide, and C-fusion polypeptides are commercially available (*e.g.*, pDEST (Invitrogen, Inc.)), and any vector can be converted into an expression vector for receiving a nucleic acid from the shuttle vector by introducing an insert having an attR site flanked by an antibiotic resistant gene for selection using the standard methods described above. Transfer of the nucleic acid from the shuttle vector is accomplished by directional recombination using Int, IHF, and Xis (LR clonase). Then the desired sequence can be transferred to an expression vector by carrying out a one hour incubation at room temperature with Int, IHF, and Xis, a ten minute incubation at 37°C with proteinase K, transforming bacteria and allowing expression for one hour, and then plating on selective media. Generally, 90% cloning efficiency is achieved by this method. Examples of expression vectors are pDEST 14 bacterial expression vector with att7 promoter, pDEST 15 bacterial expression vector with a T7 promoter and a N-terminal GST tag, pDEST 17 bacterial vector with a T7 promoter and a N-terminal polyhistidine affinity tag, and pDEST 12.2 mammalian expression vector with a CMV promoter and neo resistance gene. These expression vectors or others like them are transformed or transfected into cells for expression of the target polypeptide or polypeptide variants. These expression vectors are often transfected, for example, into murine-transformed adipocyte cell line 3T3-L1, (ATCC), human embryonic kidney cell line 293, and rat cardiomyocyte cell line H9C2.

[0265] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the claims which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0266] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.

Nucleotide and Amino Acid Sequence Embodiments

[0267] Table A includes information pertaining to the incident polymorphic variant associated with osteoarthritis identified herein. Public information pertaining to the polymorphism and the genomic sequence that includes the polymorphism are indicated. The genomic sequences identified in Table A may be accessed at the http address www.ncbi.nih.gov/entrez/query.fcgi, for example, by using the publicly available SNP reference number (*e.g.*, rs756519). The chromosome position refers to the

position of the SNP within NCBI's Genome Build 34, which may be accessed at the following http address: www.ncbi.nlm.nih.gov/mapview/map_search.cgi?chr=hum_chr.inf&query=. The "Contig Position" provided in Table A corresponds to a nucleotide position set forth in the contig sequence (see "Contig Accession No."), and designates the polymorphic site corresponding to the SNP reference number. The sequence containing the polymorphisms also may be referenced by the "Nucleotide Accession No." set forth in Table A. The "Sequence Identification" corresponds to cDNA sequence that encodes associated target polypeptides (e.g., ELP3). The position of the SNP within the cDNA sequence is provided in the "Sequence Position" column of Table A. If the SNP falls within an exon, the corresponding amino acid position (and amino acid change, if applicable) is provided as well. Also, the allelic variation at the polymorphic site and the allelic variant identified as associated with osteoarthritis is specified in Table A. All nucleotide and polypeptide sequences referenced and accessed by the parameters set forth in Table A are incorporated herein by reference.

Table A

RS_ID	Chromosome	Chrom Position	Contig Accession No. [1]	Contig Position	Nucleotide Accession No. [2]	Sequence Position	Amino Acid Position	Locus	Locus ID	A [3]	Allelic Variability	OA Assoc. Allele
rs756519	6	170707371	Hs6_7740_34:11	520890	NM_002793	intron		PSMB1	5689	F	[C/T]	T
rs1042327	6		does not map		NM_003194	exonic		TBP	6908	R	[T/C]	C
rs8770	6	170743040	Hs6_7740_34:11	556559	NM_002598	mrna-utr		PDCD2	5134	R	[C/T]	C
rs1563055	8	27976377	Hs8_23822_34:16	6328752	NM_018091	intron		ELP3	55140	F	[C/T]	T
rs912428	13	44965904	Hs13_24680_34:1	28147904	NM_015116	intron		CHDC1	23143	R	[T/C]	T
rs1888475	21	38832863	Hs21_11669_34:9	25572863	NM_004449	intron		ERG	2078	F	[A/G]	A

[1] Contig Accession Number which can be found in the NCBI Database:
http address: www.ncbi.nih.gov/entrez/query.fcgi

[2] Sequence Identification or Nucleotide Accession Number which can be found in the NCBI Database:
http address: www.ncbi.nih.gov/entrez/query.fcgi

[3] "A" column is the sequence orientation ("F" is forward, "R" is reverse).

[0268] Following are genomic nucleotide sequences for a *chrom 6* region (SEQ ID NO: 1), an *ELP3* region (SEQ ID NO: 2), a *CHDC1* region (SEQ ID NO: 3) and an *ERG* region (SEQ ID NO: 4). The following nucleotide representations are used throughout: "A" or "a" is adenosine, adenine, or adenylic acid; "C" or "c" is cytidine, cytosine, or cytidylic acid; "G" or "g" is guanosine, guanine, or guanylic acid; "T" or "t" is thymidine, thymine, or thymidylic acid; and "I" or "i" is inosine, hypoxanthine, or inosinic acid. Exons are indicated in italicized lower case type, introns are depicted in normal text lower case

type, and polymorphic sites are depicted in bold upper case type. SNPs are designated by the following convention: "R" represents A or G, "M" represents A or C; "W" represents A or T; "Y" represents C or T; "S" represents C or G; "K" represents G or T; "V" represents A, C or G; "H" represents A, C, or T; "D" represents A, G, or T; "B" represents C, G, or T; and "N" represents A, G, C, or T.

chrom 6 genomic sequence (SEQ ID NO: 1)

>6:170689051-170779900

```

1      tggcctagcc caggccatct gccacacttg catagaatgt ttttcctatg gaagccttatt
61     ttggaatata atcacagagg aaattgagat aaacaggcctt ttccttttct ggacaattctt
121    tagagataaaa atctactgca gaaggcaaga cacttaactt ctgcccagaa actgatgtca
181    cgtggttaac  gtgtggtggc cacatcaact gaagccgaca gtgacaccRa gtcagcatcc
241    tgggtgcagc  tgccaccctg tcaactctcat cacattccac cattcccagg ccacagaggc
301    aagaggatac  tctcaaaaact tagagatagc ttttagtttta ggaaaatagg acggaaaaat
361    atcttatgccc tcagaataat aggtggaatt caattgctct tcaaggcagc acaagcataa
421    tgtccaaata  ggcttacaaa ggcctttccct agattgtggg tgattgatct gtggcaggct
481    attaagggaa  gttcaggatg gccaggctat tttgaggctg atatcttgga gaacctctc
541    tctctcctgc  aaaccactcc cttggcatca ctctcagaga caagctgctg gcctgtgcag
601    cccacaccct  gactcagggt ggcatttttc tcaagccctc tccttttaag gccatcaggg
661    atctgagccc  ccatgggtag ggaacaggcc aggcagctgg agcgtacaag agtcaccact
721    acgaatagct  tgactctgaa gtttctaaaa gggaacctgg accacctgag cctcccaga
781    agacgcctca  gtcacattta ccctgctgct gccattagta tcaggcagag ccatctgtac
841    acgcgggaac  atgaagggca ctttgcaaga cttcagctgg tgggttaaat gtggctcaca
901    gctgtcgctg  aggcacactc cccaaaatag tctcaacaca gcagtaggaa agatgacaca
961    gagggtgactg agactaatag caatacagta ttactattat cccttgaata aagtatgtaa
1021   atcaagggtgt taaccatcag gagcagagaa aagagagtgt taagtatagg gcttagctg
1081   taaaaagaat  cagaagaaaa agaaaaaaa accctcagtt aaagggctag atagtctaga
1141   aaatttgttt  ggtattttata ttattaacac agagcagggt aacacgatag aatagaaacc
1201   tgttacaaac  aggcaacaac aacaacaaaa tagacaacaa ccctagaaga aaaatgcctt
1261   gaaggaataa  ctggaatact gtgagagact taccagtggc ctgtctgcgt gtaactaact
1321   cctcatcccg  acttggtcac gcaaaggaca ggtgaccata cctccaggaa ggtggaagg
1381   ggccctacat  gagtgagagg ctcaagtcc  gcccaagtgg aatctgtgat tctgtgatct
1441   aaagaaagtt  tgtggctatt tttagaaact aaagtttatc tcatattgac acaaactcaa
1501   aatcaatga  tactttgaga tatgtcatat cacaaaaaag atttctcctc tagtacatta
1561   tcattttaac  aatacaaaac ggtctgggcc aggtgctgtg gttcatgcct ataaaccag
1621   cactttgaga  ggcctgaggg ggtggtatccc tagagccag  gagtttcagg ccagcctggg
1681   caacatggtg  aaactctgtc tctacaaaa  gtacaaaaat tagccaggta tagtggcaca
1741   cacctgtagt  tccagctact aaggaggcta aggtgagagg accactgagc ccagggacgt
1801   agagactacg  gtgagacatg attataccac tgcactccag cctgggcaac agagggagac
1861   ctctctctca  aaagaaaaa  aaagaaaaa  aaacaggtct aatttgttca tctaagcaat
1921   gataagattt  atatgaacat aagttgcttt attgatgaaa aattgaacat aatctaata
1981   agcctctaaa  tttaactgcc aatttatagg aaaagacagg acagaacct  caggaatgca
2041   atcattttata tctagaatgt ggaagattct gcatgacaaa cgacttggat tcttcaacat
2101   gtaaatattca aggaaagaga gagagagaga gataaaaagg cttgtttcca gatttgaatg
2161   acatgttaat  agacatttat gagacaaatca aggaaatttg aacatggact gcatattgaa
2221   gtttgaggga  ttatagttaa tttttaaagg tacaattgtg atactgtgat tatattttta
2281   aatgcgatca  ttatctttta ggagcactaa aatattttact aataaaatta taggatttac
2341   ttcaaaataa  acaacgataa taagattacc atgaatagtg gtgggtgaaa tatacaaggg
2401   gcttatttga  catttccata ataaaaaaca cgaatgaata acaaagcatt atagaaattc
2461   aagtaaaaaa  tagctcgtgc ttagtaataa atcaaagcgt gctggacact ttaataataa
2521   aacaagata  agatcttgcc ctctagaaaa cagctctaata tcaggacaga cttgtcaca
2581   tcagatggaa  gagtcagaat gagatgtgct gagtagaagg tagaacggtc gctagcagag
2641   ggtgggaagg  tggtagtgt  gggggagagg agaaagagag gttgattaat gggtagaaac
2701   agacagttag  atggaaggag tgagatctgt tgttcgacag tagagtaggg tgactacagt
2761   taacaacaat  atttgcata  ttcaaaatag ctagaagaca ggacttgaa  tgttcccaac
2821   acatagaagg  gtgatgaact caaggtgctg aacaccccaa ataccctgac ttgattcata
2881   tgcaataaca  aaatatcaca tgtaccctgt aaatatgtac aaatactaca tatcaattta
2941   aaaattttca  cacaagaatg agatgtgctg atgcatgtgg aggccttagc tccagactgg
3001   gtgtaaactt  caagccactg aagatcttat ttccaaggtt tttctacttt gaaattccaa

```

3061	acttatttttt	ctagcagatt	tataagggac	acgggacaac	ataaacttgt	taaagtgaca
3121	agagaaagta	aatatgccct	taaattttata	ccaaatcttc	tgacaagtct	tgactgataa
3181	ttgttttcctt	caaattttgtg	aaaaacatga	gagaaaaacgt	gtttgtatct	cattttttaag
3241	tgtggacact	tggtcattgct	cacggcttcc	aacggaataa	aatagggctt	agttgttttc
3301	cattagcttt	ttcctttgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	tggtgtgtgt
3361	acataattggc	tctaagcatt	tattagccca	ataattttta	gtgaaactct	ccacttctca
3421	atatcttttg	ccgtattaga	tttttgtaac	atgtgctaaa	ggttaaaaaca	ccttttcccc
3481	ttcatgaagc	ctgagagaag	tcggttttct	gggtttttctc	aagacaaccc	agaggttttg
3541	tatacgtctg	tctaaaaagt	ctcagatttt	tcttgctaata	tgtgcacctt	cataatcaag
3601	cagacaaatc	agaatattat	tttggtgagg	ccatcatcta	aactaacagt	ccttatgtac
3661	agaagcagca	ctgaccgggt	tcatactcct	cagttagcaa	gtcaacatct	tcctctggcc
3721	agcaacccat	cccagaatat	ctgtggctta	tttaacttta	tgaaaaacaac	agtcttcatt
3781	atttactaat	taggagatga	tcagatgtat	ctattgatag	caacaggcta	tttaaaagtg
3841	aaataatcta	tcaaacagat	tttttatcaa	ctcaaagt	ccagtttagat	atttttcatt
3901	aaattgattg	ctagattgca	gccacaatca	aacttaagta	ttataagaag	tttgggttgg
3961	cttttaaaat	catgcaaaaa	ttcaaggggt	gctattaaat	atagaattcc	aaatgtataa
4021	agtcttgttc	taaaatgggtc	aataaaatga	accagtcac	tggttcattt	atagggggccc
4081	agtcactac	attaatttgg	atgttcttcg	tctgcatttt	catgttttca	cacacctgca
4141	gtcattgtgt	acaattctgg	actcttgatt	ttcatttagc	attctgcgtt	aatattctga
4201	catgttgcgt	ccaggctctt	atggccacca	cttttaattg	ccagaaaata	tttcatgtag
4261	tgaacatttc	ataaattact	taaccagtgc	ccaattattg	gttattttaa	gaatttataa
4321	ctaaaaacacc	actcttgtga	caaagtctcg	atgtatccag	atgtactcat	gccaagtcgt
4381	ccatcaatag	ttctgctaaa	ccctgtcaga	gcccttttct	gaagggaacc	aggaaacatc
4441	tcacaacaag	aagcttaagg	cctctacaaa	atgacttcag	ggttaggatt	tcaatttcac
4501	tctgaggcac	atacaggaac	cactaggtta	tttggcttag	aatggaggga	agtgctcatt
4561	tgtttctgtt	cggcctgcag	gaagctcctt	cccaggccct	gcattgcca	ttgaacaatt
4621	gaacaattct	cattgttcaa	ttcccaccta	tgagtggaga	catgcattgt	ttggtttctt
4681	gtccttgcca	tagtttactg	agaatgatgg	tttccagctt	catccatgtc	cctacgaagg
4741	acatgaactc	atcatttttt	atggttgcac	agtattccat	gggtgatatt	tgccacattt
4801	tcttaattcca	gtctatcact	gatggacatt	tggtttgggt	ccaagtcttt	gctattgtga
4861	atagaccgcg	aataaacata	cgtgtgcattg	gtcttttata	gtgcctagaa	cataatagga
4921	gctccagaaa	tactgttgaa	aaaatgaata	aattgagcac	actaagtgtc	tgaataaaat
4981	accctgacca	tacccttaaa	taaacaacat	aaataagcaa	atttcaattt	ctcggaagag
5041	ttatatattta	gtgtccaatg	ctcttgttat	gcagtaattg	catctttgat	tattcatatt
5101	cgttagagct	tccagaaagg	agtattgcaa	atcacacggg	cctctgactc	tcattgaccaa
5161	atccccctgt	cactgtcctt	gttctctgat	ccttccctctg	ggccctcgga	aatgttggtc
5221	tccatcaaat	tgctgtaaac	agttttcaga	aaaagttctc	tttgggaagt	ttcaagagaa
5281	aaccacaaat	ttcctggaaa	tgcttcatgc	ttcatgacat	ttaaggcttt	cagagccttg
5341	aacttcagga	gcaagatagc	tggttaggtct	tctggaggctc	ttgcctaacc	tgagaaggctc
5401	ccagagaatt	tgtgcataga	acctcccagg	aagcagtaag	acaggctggt	gcagctccac
5461	aggatgagaa	aggtaagagc	ttaccaatgt	ctccagtttt	gtaaaatgaat	gttcttagca
5521	tttttgggtga	ggagaaaaaa	atatcaaaacc	catcacagac	agatcagcag	tctcttgacc
5581	taggtcactc	aaggggttct	caacaatcaa	ttctaaaatg	caagtgaata	atgtcaatat
5641	aaggctagac	acaggggctc	atgcctgtaa	ttccaacact	ttggaaggcc	aaggcagacc
5701	tagaactttg	taaaatctgag	attttgtctc	atgactttgt	ggatacttct	tactaatacc
5761	acctacacaa	aatcttcaac	caagaatctt	aattgcattt	atcaacttgt	gtccttagat
5821	cataacattt	agattcattt	gaaagaaatt	tattgaagta	aaagaaataa	gaagaacttg
5881	atagcacaaa	acaaagaaaa	cccagaaaga	gagagagaga	gagaagcaga	aaaactaaac
5941	actgaaaccc	agagagagag	agagaagtag	aaaaactaaa	cactgagctc	taatctgttc
6001	tgccactagt	cagcagcctg	cacagagcac	tttaatgtat	tctctgtcca	ttagtttcct
6061	tgtttatgaa	agtatatagc	tccaaaaaaa	ttctgtgttc	tgatttttgt	ctctccaaac
6121	cacaaccagt	cccctgtctc	cttctctctc	cctctctctc	cttctctctc	tttctctctc
6181	tctccctgtc	tctctctctc	tctctctctc	tcccccttct	cctctcccc	accacccac
6241	ctcccaaccc	ctgcatacac	ctaggacacc	tccagcatag	gttactacca	attttgcaca
6301	cctccagcaY	aggttactac	caattttgtc	gggtttatct	ccttctgtcc	tttccgcttt
6361	gatctgagga	atagctgaga	tttaggacag	caacaagggtg	tacctctctc	caggttataa
6421	aacaggatta	atgattaggc	tcaaggcccc	ttcctagtca	ctcagtaaa	tctgtgcaact
6481	ggaaaactgt	ggtagcagtt	ttctgagcat	tagaaaaactg	tggttctcac	agaggctgga
6541	tgagtcaaca	ctgccatctg	gcggcctctt	ggagggtgat	gtggagcctg	gctttcattg
6601	agaatgaag	gtcccttgat	ttcctgaccc	tcagccaact	ctccagcagc	ttctcactgc
6661	agaagaggtc	agggcattgg	tcagcttgag	gacaaaagtg	gaggatcaca	ctttcgatca
6721	cctatacttt	ctaacaatca	gccctgtgga	catctgctca	cgccccatgc	tgtttttaaa
6781	atattttccc	attatagaaa	tatttttcaa	agatgattca	caagctccag	gagccattca
6841	gacaagggag	agcaaatgg	cagctaaact	cattcaagag	tgagcagat	gcacatgaag

6901	ctctgtctcg	cggaggaggc	acaagacacc	gagcctggct	gggagggtgc	tcatgacaag
6961	agtggggcca	caggcctctc	ctttcattgg	acatagtggc	catacaaagt	ggctgccttg
7021	aatgcaccc	acatacacta	tcggtctctc	cggtctaggg	agaaaactag	cacagtatgt
7081	caacagataa	atcagcgcaa	tcctgtgggt	gaagcagtg	accatactc	ttcattctgc
7141	tgagcggaa	tcaaggggtga	gacacttgta	ttcctaactc	cacttgccct	caggctcctc
7201	caacattcca	aattgcccac	ggacagagct	actaccctc	catagaagtg	acaacttgaa
7261	ataaaagatt	tgaagctcct	tcccacgttt	agaccaggc	ctgctgctag	gaactccaga
7321	ggagggatgg	aaagaagttt	gcactgctca	cagatttaac	gtttctctca	caaccagaag
7381	taggcagcag	gatggatttt	caatcaaac	taacaaacgg	atcactcctg	ggctccttaa
7441	acaagttcca	tgtctcttta	taggttttag	gtgcctccta	tgtgttcaga	cttcgtataa
7501	tgatatata	agaaaagtaa	atgggagggg	caatatTTaa	gaaataacag	caacaatttg
7561	cattcatatt	taatataaac	tacaagtatc	aagttaatag	acgctcagaa	gaaattcata
7621	aatctagaa	aagttgggaa	aatctcaata	gagaaaataa	aacttgatct	ggactttaag
7681	agtatgattt	ataaagttca	aggggacagt	acacaaagtc	agacaatgtt	agctgaatca
7741	ctgaagaatg	agaatatcac	accacttaa	caaagcagcc	cagctagtta	aagaagtatg
7801	ttgaagaggt	agggagttgg	gtgggggagg	cgaggatggc	gtggagggat	ctgaaaacta
7861	tcaaccttaa	ataatgagat	ttagaaaata	tgattaagta	gagggttaat	tcgagtgcag
7921	agcttgagaa	tggccacctg	gaaacactga	ctctaaacca	gtaaggttaa	tgtttcaaag
7981	tggagaagtt	aaggtttcac	ttagaaattt	tagcaggatc	acattttcca	tacaagacca
8041	gtgcattctgc	cacagcaatt	tgattggtta	tagattgctg	ctcattccaa	gattacttta
8101	ttactctgtg	cggaggagta	gtgattttag	gggtcttatg	tctgggtcct	ttttgtcttg
8161	tttacagggg	aaaaggcaga	agttgcgcct	gcctgcgcga	taactcaggc	tctgcatagc
8221	cacatgtctt	tcaaggctca	gaataatttg	aagttccaac	agctttaagt	ttgaattaat
8281	ttcacaaagc	ggagaaagac	ttggacttcg	tgagtaatg	aagaccaccg	atgggtactg
8341	aacagctctg	gtgggtttgc	actcagagaa	aggagcctaa	gatgcagaag	gtgcagttag
8401	gcagcagcca	tctgctcttg	tgctgtgtag	cagagcatga	tgggctgtca	ccgctcacgc
8461	gtgttcttta	cctgctgact	cacctggcgg	catgggctgc	atatggggtc	ccagctcctt
8521	aagctcctgt	caggcccttc	tgcaacttct	cccaagctct	tgggccaggt	gcctgtctag
8581	ccatgaaaaa	ggaggccagt	acctgatcac	taagtgaag	ttctaaggta	gtgggactgc
8641	cacaggtgtc	cccattggtc	ccaggtcaca	atccagtctg	ttgaccctcc	tccttttgca
8701	ccatgcgccc	tttgacagcc	tgctgtatgg	gttttaggct	cctagcacca	aacacaaaca
8761	ggcttatata	gatcatctaa	tcagctcttc	ttatgtgagg	ccgaactctg	taataaatct
8821	ttttatgtct	cctagggctt	ctctgattga	acctgtctga	tgaggagtta	aagaagttct
8881	caaattgtgt	atgcataaga	atcacctggg	atttctagct	cagagactgg	gacaataggt
8941	ctattgaacc	taggagttca	ttttgaacaa	gtgctctacg	taattctgat	acagggaatc
9001	ttccagttac	agtttgaaat	tcacaaatac	aaggaaatgag	agacctagaa	tcgaaaagca
9061	tgttaatata	cttttggggc	attaagtgtc	cacccaggta	acaggtacag	aaaggcagaa
9121	aaaggaaacc	tatcagggtta	ataattatgc	cttttttttt	tttttttttt	gagacagggt
9181	cttgctctat	cacccaggct	ggagtgcagt	ggcacaacca	cagctcactg	cagccttgac
9241	ctctcgggct	caagtgattc	ttccagttca	gcctcttgag	taactacgac	tacaggcatg
9301	tgccaccatc	cctggctcat	ttttgttaga	gagggggttt	ggccatgctg	cacaggctgg
9361	tctcaaattc	ctgggctcat	gtgatttccc	cacctcagcc	tcctaaagta	ttaggattac
9421	aggcgtaagc	cactgtgctt	ggccaagaca	ttttggttaag	aaatattatt	ttcctactaa
9481	attgtctaca	ttccccttgg	gtaggcttgc	aaagtactg	tgactacagc	aggagctatt
9541	gtcgtcaggg	aaatatggag	acacgagtgg	tacctggcag	tcacgggctc	agttttgttc
9601	ttaacctcca	agtcagcaca	gccccactga	gcagactgcc	ggaaaagtatt	tatgccatct
9661	gtcggataat	taagacaaat	ccaaacatct	acgtgcattc	tgtgtgtata	aatggagtca
9721	tggccaacct	ctcaagcagt	tttccatcaa	tcacttgtaa	tattaccaga	tacttccaat
9781	ccccttgtag	gcagtaatga	agagaaagtc	tctacatcag	cagcttctca	ctggaatctt
9841	caagcctaac	tcttaagaat	gaatcatact	tctaagtgtt	atcaaaggtt	ttcttttcac
9901	acaagggaatg	tgccatgaatg	atctataaaa	ctacccaaat	aactagtatg	cagggttagct
9961	ctaggagcca	tatgccaaac	acattctgga	ctaagtctgt	ttcactccaa	aactaagacc
10021	atgttcaaa	tctacagcaa	tactggatta	aaacaatgtt	agctgaaggt	tgtaaaaaaa
10081	catcatggaa	cactaattat	gtgccatcaa	tcattactca	atttgtagaa	cttcctataa
10141	caccttggga	ggtgagaacc	ttatagaaaa	tcctagatac	acaactttat	ataaacagaa
10201	aattccaaaa	tgggaaatcc	tttagaataa	gaatatacct	agtgatatata	aagaagaaaa
10261	aatgtagatt	tgaatttggg	atactctaca	gttttttagtt	gggacattgg	gatagttaat
10321	tttaggtgtc	aacttgactg	gattaaagga	tatcaagaca	gcaggtaaag	cattacttct
10381	gggtttgtct	gtgaggatgt	ttccggagac	tggcctagga	gtttgtggac	tgagtgcaga
10441	tcagctgtgt	tgagcaggca	caatccaact	gaatggagcc	cagatagaac	caaaaggcaa
10501	aggcaataaa	ttcactgtct	ttctcctcga	gccagggact	tgattagcaa	cccccacac
10561	ccaactcctt	attttcagga	cttcaagggt	tataccatca	gcttccctgg	ttgtgaggcc
10621	ttcagacttg	gactgagcca	tgctgccagt	ttctctggct	ccccagcttg	cagacagcca
10681	ggcatgggaa	ttctcagttc	caataactga	gtgagccaat	ttctttaata	aaccttttct

10741	catctctctc	tatatacaaa	catatcctat	tggttctgag	tcttttgaga	atcctcatatc
10801	agacactgtt	aagagatgaa	ccagggtgta	ccactgtgtc	cccaatgcct	gggatagtga
10861	gtgacagtat	tgatgtttgt	gcaataaatg	tttacaatgt	gtatctacta	gcagctttcg
10921	ataatttgaa	ttattttataa	aaggcaaaat	aaacaaaagt	ggtatctaag	attaaccact
10981	aacttagtaa	aattccattt	ataccacagg	ttgttctggt	acaaagcata	caataactcc
11041	catgtttacat	ataatttgac	ttgttcaata	acaagaaaat	tactcaaata	ttataatgga
11101	gttaaaatgt	atttaataga	tctaaaaaat	tttggctcta	tcaaaactta	ggtctcaatt
11161	ataacattac	taaagtgcct	gtaaaaatta	tcccagattt	ctggctacag	cttgataaaa
11221	ggcctatttt	ttttttttta	cagttttatt	ttggttgaag	atttttctcc	aaaatagcta
11281	ttccagatgg	cttaacagtc	ccagaagtga	aaaatcttaa	gatatttcat	ttataaccat
11341	tagagtctta	ataaaacctt	agtaaatact	ctcccttctg	gatgggttaa	aggtcccttc
11401	aggcaagctg	gctgcgtaaa	caccagagcc	ctcttcataa	gataagtttg	ctcaaaaggt
11461	caacttttac	agagaaattc	ctacctcatt	agtgttaagg	aaaattaaca	tgacctgttt
11521	tcattactta	aaatgcaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaatctcag	aaaaatagaa
11581	aagggtggga	aaaatgaaga	aaattaaaag	aaattctact	tcctatatct	gtcctgtcta
11641	gagaagacaa	gttatagcaa	aatgagtact	tcagggttct	cttttaataa	acaaaacctt
11701	ccaaggtaca	gttccaaagt	acaaaatcaa	ccagggtctga	actgattggt	gataagagca
11761	cacagatcag	tccttcctta	aggaacaggt	ttcctccctg	atgccctctt	tggtcactat
11821	gcagatccgg	agtgcgtccM	cagtgtacac	atctctctca	gccgcagaaW	tgaagacatc
11881	tttaccagc	cgcattggctc	tgtccaagga	cagcggaaca	tgctccacat	tctgcattgt
11941	cttaaaacca	acctgggtggg	acatgaaact	tggtgtgagat	gtcatgtgac	agtgagcaca
12001	aaggaaaggt	tcaagtttct	tccaataccc	tcatgtgtac	gaggatttgt	gatgaaaata
12061	ctaYacaaaa	ttaaaatgtc	cttcagactg	aataaagacc	cttcagagcc	tggtatgaatt
12121	acacacatca	gaaatctgac	aacgattccc	tggcaaagct	cccactcttg	agaagcacat
12181	gaaccggaga	ggtggacgta	tgcaagcaaa	gtgatcagag	aattgctata	actaagaaga
12241	aaataagagc	caggagagcc	ccaacagatg	tccactaagt	ctcctagaag	aagacaggct
12301	tcagcagaaa	ggacttttaa	gtagggtgagc	gctgaaggct	acactgaaga	gcacagaaca
12361	ttccagactg	aagatgtggg	acggagttag	gaaaagcaca	tcttttctga	gaaaagcaga
12421	gttaggggtct	gggtgtagtgt	gggtgacagg	agatgaggtc	aataagcaat	aagcagagga
12481	gtgcctatat	atgacatcct	tcagaaaact	gagcaggagg	aattagggaa	gtaaacttgc
12541	aggcagaag	gatgttagga	aataaatcca	ataatctaata	aatgtagggtc	tgaaatttagc
12601	cagcagtaat	cacaaagagg	gacgaattag	aaatacaggc	ggcctttgaa	taacaccggt
12661	ttcaaccact	caggcccact	gacacataaa	ttttcagtaa	acacactgag	aagtttttgg
12721	agactcgcaa	cactttgaaa	aactcagatg	agccaagtag	cctaaaaata	ccttaaaaaat
12781	taagagaaaag	gtatgccaca	aatgtataaa	ctatatgtag	atactagcta	ttttatcatt
12841	tactaccata	aaatacacat	aaactgatta	taaaaagtta	aagtgatctc	tcaatgttct
12901	cctttgtttt	tcattgtgtt	tagtgcgatt	attataaacc	ttaaataaca	agtggcccaa
12961	acaaagcact	attagtgtatg	ctggaactgc	tccccaagaa	gcagggaag	ttatgacatc
13021	acaagaaaaa	gctgaattgc	ttgatattga	ctgtagattg	aggtctgcag	ctgtggttgt
13081	ccaccactcc	agacagatga	ttctccttgt	aaacagatga	tgtaaaactta	cagtatcaac
13141	aaatacagta	ttgtaaatgt	attttctctt	cctataattt	tcttaataac	attctctttc
13201	ctgtagttta	tttatcataa	gaatacacga	cataatacat	atgtaaaata	cgtgataatc
13261	cacagtttac	attattggca	aggcttccag	tcaacagtag	gcttatttgt	aagttttctt
13321	ggagtcgaaa	gttatatgtg	gattttccag	tgctcgggga	actggcgccc	caatccccgt
13381	gttgttcaag	gStcaactgt	atgcttggga	gttcaagtac	agtattgggt	aacacagaaa
13441	gggggggttaa	gagggaaaaa	ggcattttta	tgatcttcag	gtaaagacag	caagagaatt
13501	agccactaca	gacataaaaag	agttggaagc	catgagaata	caggaaaaca	atcaaaaacca
13561	tgggagtagg	taaggaccca	caagagagta	ggcagaacac	agagagagac	catgaggttt
13621	ctgtaggaca	cgtctgccaa	actcaaacag	taatttataa	gaaagtcttt	atatcttggtc
13681	tctcctaagc	taatgaagag	gagaacttgg	aggaaggcac	ttcaactgaa	caccatgaca
13741	agttggacat	agtatcatta	cctggacaca	gtatcattac	ctggttgtca	agcaggggct
13801	gtagcatggc	acttgcctgag	cctccagcct	tgaaggagtc	tctctggtaa	gaccctactg
13861	gatcaaagct	gtatacagcc	ccctttcctt	aaagaagaaa	acagtattca	taaggatggt
13921	atccaatcac	aatcccaaatt	catcgcaaaa	ataaatcagg	tcaaaacgtg	acacaaaacg
13981	gactatccac	ttggctaaaa	cttcagggaa	cagttggaac	acagccactg	ttgcactggt
14041	gacagctgca	gcacacacac	tgttgacacat	gtgacagcca	gaaccacgac	accattgac
14101	aggtgccacc	cagcagacag	ccccaatccc	cgtgttggtg	tggtgttcaa	gggtaaaactg
14161	taggcttggg	agttaaaaata	caatattgag	taacacagaa	aggggtgggt	agagggagaa
14221	aggcattttt	atgatcttca	ggtaaaagatg	acaagagaat	taaccactgc	agatgtaaaa
14281	cagttggaag	ccatgagaat	acacaaaaat	aatccaagcc	atggtagtag	gtaaggacc
14341	gtaagagagt	agggagaaca	cagagagaga	tggtgaggtt	tctgtaggac	acatcagcca
14401	aactcaaaca	gtcattttgta	agaaagtctt	tatttcttgg	ctctcctaag	ctaacgaaga
14461	ggagaacttg	aaggaaagca	cgacttatgt	atatagacag	gtccaaggga	cacatccgca
14521	agacacagta	aaacgaagag	aattctgtaa	cctgtaaaaat	acgatactgc	ttttgttttaa

14581	aagaataaga	aaactatttg	tacaagcaca	aaaaagtctg	gaaagataag	taactgtggc
14641	tggtcatctt	tgagctctga	attagaggca	agaaaagaag	ggaattaatt	catgtgcttc
14701	tatcttattt	gagtttggtt	caaagaacac	aggtgattct	acaatcgtta	aaaaggcctg
14761	ttaaaaacaa	ccgtgttctg	ggtatgagaa	atcacgagtc	atttacaagg	tatgattttt
14821	taaaaacata	actgatttag	aaaacacatt	atataccaga	actcagtacc	ctgggagatg
14881	tgatataggg	acaatctcct	tactaaccat	gtacaactgc	cactgttagc	tttgtatgtt
14941	ttctactata	gaactgcgta	aacttgGCCa	ggcgcggtgg	ctcacgcctg	taatcccgac
15001	actctgggag	gctgaggtgg	gccgatcacg	aggtcaggag	ttcgagatca	gcctggccaa
15061	cacagtgaaa	ccccatctct	actaaaaata	caaaaattag	ctgagcatgg	tggcacgtgc
15121	ctctattccc	aactacttgg	gaggctgagg	caggagaatc	gcttgaaccc	aggaggcgga
15181	ggttggtggg	agccgagatc	atgccactgc	actccagcct	gggcaacaga	gcgagacttg
15241	gtctcaaaaa	aagaaaaagg	attgtgtaaa	ctgaattctc	tgggtggagaa	ggcaccatct
15301	gaggtcacct	aagaaactgg	gaagtatttt	tgctggaagc	aagcttgggg	caagacatga
15361	gatcacagca	ggcaatccac	taacctcaat	atacaaaaaa	agaaagagat	gtaaggtgta
15421	aaggcagaaa	aatgattgta	tctctcatga	ggatgacagg	ccacaagtaa	actgaggccc
15481	aaaagcttta	ccaagggatg	gaaccaatct	cagaaaacca	cacacaacaa	atttaaaaaa
15541	ataaaccaaa	gaaaggggtg	gctatgtcca	agaacacagc	tttaagacc	ccctagaaga
15601	agcaagagag	acactagtaa	aatcagtata	aaacagcaaa	aatcaataaa	taaaaacaaa
15661	aataaatgta	gaatggcatg	aatcttgatt	acaaagtgtc	tttctctatc	aaagtaaaat
15721	ccaaaaagca	caatcatttc	aatgtaaaaa	ataaaatcag	acaagtctta	aaagaaatac
15781	agcttaagctt	tgatgggag	aagtttttct	aacaaggcaa	gaaactgaat	gtcataaaga
15841	aatgattgac	agattagact	gtatgaaaca	cagcaatttc	tgtgtagcaa	aaggcaccac
15901	ccaaaagaga	caaataacac	tcttgggagt	atctataata	tcacagacat	ttgacacagt
15961	cgaatgtagt	tcttgaanaa	cagtaagaag	aaaactaacc	tgcaanaaag	gaaatgggca
16021	acagtttaata	tatgaatcag	tctccaaaga	ggaaatttat	aaagtcaata	aacatatgaa
16081	aacatgccca	agaatatata	taaatgcaaa	taataaactg	ttcaataggc	aaaaagatgt
16141	ttcaaaatgt	gttaattcta	atactgggaa	ggttatataa	aatcgggtac	tcagggtact
16201	ggcgtatcct	tcagaaagag	tatactcaac	aaaatacaaa	aaaaaattta	tctgacccaa
16261	tacttcaact	actcgaagtt	tattctagca	agacaagcat	aattggggaa	gtacacaaaa
16321	atgtatgtac	aaggatgttc	atcatagtgt	ttaYattaag	aaaacaggaa	gccaaaatgc
16381	cctttagagg	actggttagg	taaatgtatg	tatgtattac	agcagcaaaa	agcaaatgca
16441	aaaaaatttt	ttgttaatta	aatacaaaaa	tcattatgca	tatagcttag	atctccttga
16501	taatacataa	tgtatagtta	cagaaaaaca	gataccaaag	tcctacttac	ataaaactRc
16561	atgagtgtat	gtataaacat	gtactgattg	ttttcactac	agaaaaacta	aagccacttt
16621	taatttggga	gaaaaatttt	taaaaggcat	gagagcagct	aagagattaa	agacagagag
16681	gttagtgcag	acagcgggtg	acaccaaggg	aaagagctgt	ggaccttgct	gtcctgaagg
16741	tttcaatatg	gagcactgct	ctccaggctc	ctgcatccac	caacacagaa	gacagtaaag
16801	gaagtctcct	gagacattag	cagaggctct	ggacttaggg	gaccaggctc	tagagtggca
16861	aagggttagga	gcaaaacaca	tgaatggaa	agttgatgta	aacaaaaacac	agaccttcag
16921	ccccttctct	cctcYgaggc	ctaaaaagct	gccaaaccgg	cttctcttcc	tgtagtccca
16981	cctcacctca	ggaaaaaaga	ttaaagaaat	tcttcatctg	aagaaaaatgc	tgcttactta
17041	agcaccagct	taacggctgt	cagttgacaa	gccatgtcaa	caaggacttt	cataatcaaa
17101	ttttgggtaca	agactcttca	atgtgaatgg	acagcaaaaa	aaaaaaaaaa	aaaaaaaaaa
17161	tcagagattt	gaagggaagcc	tctaacatag	aagacagagg	ccaaaaatga	acacataaca
17221	tgaaaaaaaa	aacattctga	tgaatatagc	aaaactgcaa	agaaaagtaa	caggattttt
17281	tttaaaaaaga	caaaaaataa	gtacaagctc	ttggaatta	atgacaccag	aaataaaatgc
17341	aatcagtaaa	ggatgaaaag	ctaaaaattaa	ggaaatcttc	ataaaagtaga	acaaaaaaga
17401	cagacatgga	taacatgggg	agtcaaaaat	aaagtaccgg	gtcataatcc	aggtggttgt
17461	ggagaagggt	cagaggggta	ggggaggagg	agaggaaaga	aggtggagaa	agaggataaa
17521	acaaataaga	aactctaggc	aagggaacatc	actatacaat	ttcagaacac	cagggtaaag
17581	aaaagatcct	aaaagcaaa	tctcaaaaag	gtcccRtttt	gtttttttca	gaaatctttt
17641	aggaggtaag	cctctaaaaa	aaggaaataa	atcaagaaag	aggagaacac	aggatgtggg
17701	ggaatctcag	gatggcagtg	gtccatcctg	gagcaaRggg	aaaagggatc	cacctagtat
17761	gtgaggtctg	acaaccagat	gaattgtact	cagagaatgt	acaaggaaat	gacaaaaagt
17821	ttaaagaaat	ctgagccaat	tagagatagt	taacttaaga	aaaaaataaa	aacaggttac
17881	acaagaaaga	aaatattaat	gcattgctact	tggcgcaaca	aagatcaata	tctattctggt
17941	cataataaaa	acattaaagg	tctatgttac	caaaagtgtg	atgtgttaaa	ggaggatgct
18001	aatgaggaag	aggacacaaa	agagctaaga	accatcatct	accagaacaa	aacaccaaaa
18061	gatattctcaa	aagagtgcag	tacaaatata	gagtttttga	aatggagggt	aagtaccagg
18121	aaatacaggt	aaaagagctg	aaaactgtct	tagggcaatg	gaaaaaaaag	atagaggact
18181	gctttattat	tcacacaaat	tcctctataa	aataaaaaat	agtttaaaag	acacttaaaa
18241	gccaacccaga	cagacaagct	tagaaagaca	gttctagaga	cacctgaggt	tgtcaccaaa
18301	actagcaaca	ctgttacatg	Raaatcaaa	aacagggtct	tgaagtgaag	caaaagccaa
18361	acctgaatct	taggtctgct	gcttagtagg	tatcagcttt	cccagattag	actctctaac

18421	tctcggaaaa	actcaaagga	tgctgtctgg	ccYgctgaaa	ttcaatggca	gctcagctat
18481	ttttaatggt	gctaaaatct	gaatctatta	atctgacatt	aaagccaggc	ataatgtaag
18541	tcaaaaaaga	caccagaagc	atccagcttg	ctaagttagg	caaattaatt	tgtagctaga
18601	ctctaattac	ccatgttaaa	ggtaattcat	ttctattttt	actatttcta	cactttactt
18661	tttctcttgg	ttttcaagta	aggttgaaag	ccttaaaaag	aaattgtgtc	aacctaat
18721	gtctacgcct	aaaccatggc	aggtaaaatg	atcagaatgt	tccccaaggc	tgtgtatggg
18781	ggcggatggt	aagagtcaga	gatagcgcac	Rcaaagaaaa	cgctcatgga	acagaaagcg
18841	ataggagtca	acaaccgaga	ggcagggcgc	tgaatgaag	cttgaccctt	tcctcttaga
18901	attttccctc	acaggattta	gggtccccag	tccccctgct	ttctttctgt	gtccaactga
18961	gatgttctgt	gatgtggcca	gccgatgta	ttccccctga	ggctctgacc	caagccaagg
19021	cctgccttga	acattcccag	gcgctgacaa	tgttgtttag	gctattgctt	gaaacactga
19081	gaaattaccc	ctgttgctaa	acacacagaa	actatccctg	gccctgaacc	aaattccttc
19141	aacactcaaa	tcaactccct	caccttgga	acatcttttc	ctgctgtctg	ctgagaggac
19201	tgctgcagca	ctaaatgctt	ttgactgatt	accctgagat	ttagtgttct	ggaaccccaa
19261	ctagggttcg	ggcagtcctt	tacaggaatt	cctctgcccc	tacttttagg	gcaacttcag
19321	ccacagggtc	agccagaagg	aacacagaag	atgggcttac	tccaatgata	tcctcatcta
19381	gacacttctg	taacttctta	tctcaatgga	ctactctaaa	tgacacagat	ggactactta
19441	tccatctgtg	tcaatggact	actctaagt	acaaagatct	gtcctgggac	actgtacttg
19501	tatttgatgt	ttagatgcaa	ccaaaagcca	atccaaatga	gaacatttgg	attgcaaattg
19561	agaacaaatg	gattggaaca	tttgatgga	aaacattggt	gcccattgtg	cgtaatgcag
19621	ctctggaaaca	aaaaatgggat	ggaatattaa	gtaacaagac	ttaatatctt	cagggaaatcc
19681	tggtctgact	ctgacaaaat	cagagaggtg	tttcaaaaag	ttctgaaaac	actggagtag
19741	atgcttatta	ttaattaccc	caaagtactt	ctagaaccgt	taaagaactt	ttttgaagt
19801	tatttggttaa	taaaaaaatt	ttcaatactc	agaatcacat	ttaggtcaag	ctttctaggt
19861	ccgatttctt	tctataagat	aaaaatgcct	aattatttac	tactaactca	tgcaagttcc
19921	tgagaaaata	gcaaaataac	tggctgtcca	acattcacta	ttagagcttc	aatgtgcctt
19981	tgagggtcat	acattagctt	cagtttcttc	atcactttcR	ttatcagaaa	tgaaatgaga
20041	attgtgtcag	atgggtattta	tggttctttt	caactctaga	tggtatcaca	catctaggga
20101	ataataactc	ctccccccca	ccattttcca	gaaagaagga	attctacctt	cttcatcaag
20161	tccaccgatg	atgttgtaaa	catagtatgg	aaagaagcgc	cttgaatata	ggattgtaga
20221	cgcatttgca	gcaattggcc	ccgtatgcat	ggccttatta	ttggaatgct	tatacatctg
20281	caattattga	taaaagtcac	aggcatgtag	aggcagaggc	cattatagac	tagaacaatc
20341	aacagtataa	agtgaataaa	taaatgcata	ccacagtggg	aggggaaagc	atgcctat
20401	ccttcatgat	caaccaacac	agtattacta	acataaaat	ttaaggggaag	gaatgccaa
20461	tttgccctca	gctctctggt	ttagactcta	ttttccatac	cacccttatc	tcagggtact
20521	ctaattccaag	agctaattgt	tatgaaaggg	cacatgaagt	ggcctaagga	agatgggtt
20581	aaatccatct	ctctccctca	ctctgcata	agaccagggg	tcaacaaact	acccctgtag
20641	tttggctcag	gggccaatc	cagcctactc	cctattgcta	tgaatattgt	cogtggctgc
20701	tttttcacta	cagaagcaga	tcaaatgagg	cccataaagc	tgtaaataat	tactatctgg
20761	tcctttacag	aaaaattacg	gcaatccctg	atcaaaccta	gatgtaagtc	cctctcttgg
20821	tgagcagat	ctctgtgaga	acagaagaca	cagggcagca	cccttactct	gcatgatttt
20881	aatacacagc	agtggctcgt	atccttaaaa	catctacagt	agagaacaca	agcatccaaa
20941	acagggtcct	agcatgacaa	tacagtgact	gcaaatccta	gctttactat	acctttgcga
21001	gccctacatt	acccactggt	agagaaactt	cacaaaaatt	ataaacataa	aatgaattca
21061	agatcaccaa	tttttggttaa	ccacatagca	aacatctcta	atgagaaact	taattggccag
21121	gcatagtggc	tcacacctgt	aatcccaact	gaggcgagg	aatgtttgag	cccaggagat
21181	gaaggctaca	gtgagctaca	atggcgccac	tgtactccag	cctgtgcaac	aaagtgagac
21241	gctgtctcta	aataaataga	cagatagata	tctaactagc	accgtaaaat	taaataatga
21301	cagtcaaaa	actgaaaaat	gattagaaat	gaactacaag	taacaggtaa	gttttttata
21361	gaggaatgcc	tttccaagat	aaaaactata	aaaaattatt	catagggcgg	gcacagtggc
21421	tactgctgg	cctgtaatcc	cagcactttg	ggaggccgag	atgggtggat	cacctgaagt
21481	caggggttcg	agaccagcct	ggccaacatg	gtgaaacccc	gtctctacta	aaatacaaaa
21541	attagacagg	ggtggtgaca	ggtgcctgta	atctcagcta	ctcgggaggc	tgtggcagga
21601	taactcattg	aacccaggag	gcagaggttg	cagtgaacca	agatcatgcc	attgcactcc
21661	aScttgggca	acaagagtga	aactctgtct	caaaaaaaa	aaaaaaaatt	tcatttcatt
21721	tcatttgcga	accaagcacc	aacatttatt	tacttcacca	cctaataaag	tatgagagaa
21781	aaacttcagg	aaggggtgtg	tattatgtag	aattcaccat	tccttgtttt	attaacaata
21841	tcatgtttac	ttttgtcttc	taactcccat	ttataaaata	agcattagga	attatcatat
21901	aaatatgcc	ataaatgtaa	gtcactatgg	ctttcaatgt	aagtttctaa	aaaagcatta
21961	acaattgata	aaactagagt	gaaatccttt	ctacacattc	aaaaagcatg	tgtcagttaa
22021	ataaatatac	tatatactcc	agtaaccact	cataagtgt	aataatctca	agtggtattc
22081	atgaataggc	aaaacagaat	gtgatcacc	ctgacttctg	gcagctggct	atcaactgaa
22141	gctggtcagg	tgacagacgt	taatggtgca	tgttggtcct	cccatccaag	ctcccagctt
22201	attcattctt	tcgtatttct	actcatgcat	cctgccaaact	aagagcacag	cactgtaaa

22261	aggaaaagct	tgtcaactga	tggatgatgca	aaccaagctc	ctgtgtcaaa	tctttcctgg
22321	gctctctcct	gccacagaaa	ttgattttgct	ttgtttttga	gggaagagta	tcttcagaag
22381	aagaactcct	cccaggctcc	ttctaaaagg	gtagagaacc	cttcatatga	tacaagactt
22441	taacaaatgg	aatgtgctgc	tgtttcacat	tatgctactt	taatgaccct	gcattaatga
22501	tatcatttca	tttctagcat	acatttgtaa	ttcatctact	aacctagtga	ctctctagct
22561	atctgaccaa	cagtcattgct	gtaggcttaa	aagaatttaa	tgaattacta	ttaattttaa
22621	atagtgtaga	aatgtacaaa	agcatctacc	tttagtcttg	cttcaataat	ctttgtcagc
22681	gtaagacagt	ctccatgaaa	accgctgcat	ccaatgactg	ttttgtctgt	tctgtaaaaa
22741	gcacatttca	gaaaactgag	ctggttagat	agtagagcaa	aacatatctt	cggcaatttt
22801	aaatttaaaa	ttctgattcc	cactattccc	taaatggtaa	tgggacaaac	agtcaccctt
22861	ctggggagga	ggggtctgta	ttccatctta	ccccttattc	cagatgaatt	acaggtggat
22921	caaagacttt	aaaaatgcga	ctatgaatgg	actaaataaa	tcatgggtga	atttttttta
22981	ctatcttggg	ctgaggaaga	tacgtaggta	ggtaatttga	ctgatttttg	agacggagtt
23041	tcgctcttgt	cacccaggct	ggaatgcagt	ggcgcaatct	tggctcactg	caacttccgc
23101	ctcctggggt	caagcaattc	ccctgcctca	gcctcccaag	tatctgggat	tacaggcacc
23161	cgccaccaca	cccagctgat	tttgttattt	ttagtaSaga	tggggtttca	ccatggtggc
23221	caggctggtc	tcaaaactcct	gacctcaggt	aatccacctg	ccttggcctc	caaagtgcctg
23281	ggattacagg	tgtgaaccac	tgcacccagc	ctcgtgaagg	tagattttaa	caacacaatc
23341	caagggccac	acagaaaaat	ttcataaatt	tgactacata	aaattctaca	gggaaattta
23401	agaatagctt	tttaaaaccc	taagaaaaata	tattagtagc	atctaMcccc	aaagtcta
23461	tttaatttcta	taaaaattct	tacaaaataag	gaaagcaata	acataaagga	aagatgggta
23521	acttctatga	acagacaagt	cacagaaaaa	agagacatta	agtcaccata	aacataagag
23581	cctctatggt	actcagaaat	ccaaatctca	aaaagagact	atctcaacca	attatactgg
23641	ctaaaaatgc	ctaactaact	aaatgctggt	gaggttgtgg	aaacagtata	aactgattaa
23701	cctgttagat	gggcaatttg	gtggatttta	ttaaaaattt	aaatacacac	cactctcacc
23761	aattctattg	ctagaaatcc	atcttgttta	tatacttata	caagtatgca	aagggatgct
23821	ctagactcac	cagtgaagtac	tcattttaa	agcaccacca	aacaaaacaa	aacctaaatg
23881	tccatctgta	tggagaggat	taaaaatttg	tatagacata	caaattcaac	cgtgtatttc
23941	aaaaattaat	tcctagatct	tgatatgaat	ctataaagaa	acctaataca	tattttgcag
24001	aggccacaca	attcctcttg	gaaacttacg	aatgactcgc	acaaaccata	cctgtcttac
24061	agggctcagg	gctgcacacct	ctagggtcac	cactaacatc	acctgatcta	ctctaccat
24121	caacctcact	gcactaaggt	gaagacgtcc	aatccccctat	aaggctttac	tgcctttact
24181	gacaagattg	gttgacaata	ccaacaggga	aagacagaaa	tcttagaact	tcaagatgct
24241	tttcccaaat	atccagggtgc	atcaaaggaa	tgggagtttt	catgcaaaga	atacagtctg
24301	aaacaggttt	ggaagaggca	agcttagttc	taggtcagMc	ccacaggacg	tgggatgagg
24361	gatataata	ggcattcgtt	aatgtcgc	tgttcttatt	ctctatctct	atatctgacg
24421	tgtttcacaa	aaaaaaaaaa	aaaaagtgtc	cacttcacca	gcaaacgtaa	ctaaagcaat
24481	attttaaaga	tgagtaaaag	ctagtacaag	gatggtatcc	ataaagtgtg	tttaaaatct
24541	tattttcta	attttactact	ttcaagtgtg	acaagtgtcg	tccttgaggga	gaaaaaaagg
24601	taacacaaga	gcaccataaa	cagaaagcag	aaaggggga	tcaaaagatg	caagtgga
24661	gaaacagaac	tgggaagacg	aaaaacaaact	tcattgcttt	ttaagatgtg	ggccatccct
24721	aggagcagga	aagacaacgt	atcttttctt	ctgtacctac	ttcctacaat	acaaggaggg
24781	tccatccaaa	ggacctaaac	ctcgtaaagtc	ccattcctat	tacaattcaa	gtttaattaa
24841	cccaggaatt	catgaccatt	tataagcatt	tccaaaactg	gtaaatacag	accactgcca
24901	atctgcagta	tgtattcagt	atttatgcag	gctttttgtt	tttttaagtt	tggcctttat
24961	tttctatgtt	taggaaaaac	atagctagcc	tattaaaact	gagctgtgga	cataattgct
25021	taggatattt	ctaaaacgaa	tgtttcagggt	aaaaaaaaaa	agtgtgggga	ggcagattta
25081	aaaaaaatat	catttaaatg	attaatgggtg	ctgtggtttg	aatattccct	tcaaaactca
25141	tgttgaaatt	taatttgccat	tgtgatggta	ctgggagttg	ggaccagggtg	tttaggtcct
25201	agggctcagc	tttcatgaat	ggacattatc	acagcagtgg	gttcgcttgc	tcttctttct
25261	ttctctggcc	ttccaccatg	ttaagacaca	gcaggaaaatt	tttcatggtg	aaatgctggg
25321	gtgaacacat	ttaggttacc	gaaagcactt	ttggtaccct	gaatacagca	aatattatta
25381	agactgcaca	ttaaattatt	aggaaacatt	aacttagaaa	atggttttct	aataaaaaatg
25441	ctcccaacag	caacttaaaa	actcatgaaa	caaactcattt	agaagtagaa	actctcacia
25501	catYaaatca	ttacaaaggc	attgtgaaat	gtctttagaa	atattttactt	actatttcta
25561	acatttggtg	ctatcccgcg	tatgaattga	aaacccttca	ctcaatcgag	tatcagaagc
25621	aacaattgca	aaatcttctc	cagcaattgc	cagtatagta	ctgaggaaaa	aagaaaaaaa
25681	ttaattctcc	agggtggtaa	tcctatccct	acaaatagaa	gaatgctcca	tagtacataa
25741	tgggataaaa	tactctagat	gtcaacaaaa	acatgattca	aatgggaaga	ggaaagatga
25801	cggggaagag	aatgaacgcc	tggctacagag	ttgtctggga	aaaaaaaaatt	attaataagc
25861	ctaaatcagg	gcaaagtctc	cttggcagga	gttaacagaa	aagccaatga	attatcatca
25921	ccaacacatt	aaatacttac	tcgcgcaagg	tactactaat	acagaacaac	taaataccac
25981	atctgtgccc	ttgaggatca	ggtatagaca	gtggtactac	aacgcaagct	ctatgagttt
26041	agagaagatg	agattttttt	ttcttgcttc	atttctttat	atccaagtcc	ttatataacg

26101	cctatataat	gcttatttct	ttatacccaa	tcccttatat	aatgacaaat	agatggacaa
26161	acagtaaatt	tttccctctg	tggctgtaca	atttgacagc	ttatcaaaga	gacttacagt
26221	agaattccaa	aagcagaccg	cctgggttct	aattctggct	ttcccgtttc	gcagatatga
26281	gactgtgggt	aagttacttc	tcaaagcgtc	aatttcatca	tatatacaac	agagatcact
26341	gcagttgcta	cctcattagg	gtgttcaaag	gatcaaatat	gtaagccctt	atagcagtcc
26401	ctgacatgta	actggctctc	tagtaagtgt	tagctataag	tgctatggca	ctggagtatg
26461	actaagcacc	tggtgctctg	aattacatga	gacagagacc	cactcttgct	acttactagg
26521	tatgtgatct	tggaacaaatc	ctccaaatgc	aagttgatga	taacagtacc	tgtgtcacaa
26581	ggtgtgtata	tatatttggg	tgtgtatat	ttaatgtaca	aggcttgact	gataactata
26641	accactgctt	caatgcaata	gtggaaatta	aaggcatggt	gcctcacaga	cgtaagcact
26701	caggaaactt	aagccactat	ttttactgag	gagggatttg	tgctaaagct	ctcaagaaga
26761	aaaggatggc	attccaggta	atataaacag	caagcaatgg	caaacaggta	attattcaaa
26821	tagtacatac	attcaagcaa	ctcattcagg	cagccctttt	tgcataagca	catgtagtga
26881	cgtaaagggt	tatgtgatgg	acagggttcc	tactgtagaa	aatcccaa	gccaagctaa
26941	agattttgga	atttttagcaa	gaaatcatga	aggatttctg	agcaagaatg	atctgtagt
27001	gtaactactc	aagaggctga	ggtgggagga	ctgcttgagc	ccagggtgtc	aaggctgcag
27061	tgagctatga	tcgtgcctgg	gcatttagagt	gagacctggt	ctttaaaaaa	ggaatgcaag
27121	agagagaaaa	gttccattta	caaagtgggg	ttttaggaa	actgctctga	caaYaacata
27181	gtatgtgaaa	tggtgacagaa	acactgttct	aataactacta	atgcaatagt	aaggtagcag
27241	ggtgaacagt	aaatccaaaa	tcatacaaaa	cacacaaaa	agacaaattt	ttatatctac
27301	gcaaatgttt	taggaactgg	gaaaaccaat	tatgacatcc	aagattttaga	acttatagta
27361	acagaatgat	ggcataatta	taagtatttt	aaaggagagg	aggccgggca	cggtggctca
27421	cacctgtaat	cccaacactt	tgggaggctg	aggggggggg	gggggggtcaa	ttgcctgaga
27481	tcaggagtct	gagaccagcc	tggccaacat	ggtgaaaccc	atctctacta	aaaatacaaa
27541	aattagccag	gcgtgggtggc	aggcacctgt	aatcccagct	actcgggagg	ctgaggcaga
27601	aatgcgtgaa	cccaggagtt	ggaggtgtga	gtgagctgag	atcgcaccgc	tgcaactcgg
27661	cctgggtgac	agagtggagac	tctgtctcaa	aaaataagaa	gaaaggagaa	gaggagatga
27721	aggggaataa	ttagcttgct	ttttgttttg	ctagctgtct	tgagttgccc	tgagagcaga
27781	aaaaccagtt	aaaaatgttt	tactgaagaa	gccgaatcga	gggactcatg	agaggcagaa
27841	ctggaaaaac	agatttgga	gtaatccctc	cagcaatgag	acatgaaaga	gtgctgagcg
27901	ataaacaagg	cgctaataga	cttaactaca	tttaaagaca	gagtaggaaa	agtaaatgag
27961	gcctcatttt	gcggaagcga	aggctgcctg	agagccagct	gcagtaaWca	ctaaagaaaa
28021	agaacaatga	ctgagaaaaa	gtaatcagaa	agatctaagt	aatttttagg	gcagtaatgg
28081	cttaaaactg	attacaaggg	attaaaaagt	gagtaacgag	tagggcatat	tgaacactga
28141	aaattcttat	ttatagagaa	tagccttagc	aaacgggtcc	aataaccctc	cctacaatat
28201	acaacttaat	tagtcatcac	aggaagtgtt	aaggtgtata	atggaaaaagc	atccataaac
28261	tcagtgggtga	aatagctatg	aattaagtcc	tggctcaact	tcacaccagc	tctctgaccc
28321	tgacagttta	acgtctaata	taaccctagg	atgctaatat	catctaacat	tcacttttca
28381	tgaggattaa	ataagatgac	agcttgcaat	ttacaaaatg	catctctctt	gattctcacc
28441	aaaaactatg	aagctactaa	ggaagataag	gaaatttagg	ttcaagaagt	tcagaagtac
28501	ccaaagtgtc	ctttagtggc	agaaccaagg	ctaaaatcag	acttttcgtta	tctttctaac
28561	acactcccaa	aatgtgcatt	tatatttcaa	atttatgagg	aaccaattaa	catttttgct
28621	ttgtttttta	aattttattt	tgtagagatg	gggtcttgct	atgctgcgca	ggctgggtct
28681	caactcctgg	cctcaagcga	tgatcctcct	gccttggtct	cccaaagtcc	tgggattaca
28741	ggtgcgagcc	acactgccc	gccaatat	tctgttttaa	gaaccatcgg	ttcgttcaaa
28801	ttgcgtgtgt	atattttta	gtacaaggct	tgattggtaa	ctataaccac	tgtttcaatt
28861	tacagctctt	ccctgtcaag	agtcttaaac	agagcatctt	tctataaccc	taaatctctg
28921	gcgtgccacc	acggaaaatt	atactactca	agataaagct	ggtaatttaa	ataaaaaacca
28981	aaacttgaac	ataacataca	agaacacaca	tactaaaagg	tccatcttct	gagtattttg
29041	ttttcctgaa	cttaagctaa	acgttataaaa	aaaaaagcac	ttatctatga	aactaagttt
29101	gctcagccaa	tcccaccttc	tatttgaaat	aaaacaaaat	gattaaaactg	ctacaattag
29161	aaataacaga	aatcaggcgg	ctacaattag	acatctcggc	taccaaccca	gctatgcac
29221	taacaacaca	gaccaaacaa	ccctaacttt	taagtttcag	acgctaacc	tctaccctcg
29281	ccggctggca	taaRaaacgt	gtacatgagg	tccagtttta	atggtcttcc	acagagcaga
29341	ggctatgttt	caatttctac	tttactgtct	tacagcagca	aggagcacgg	agtggtcggtc
29401	gcataaaaa	ctcaaatgac	atgactgtaa	tgggaaaccc	taaaaaccaa	ggctgtatcg
29461	caatcaccaa	gtaaaactga	gcaaagcgag	cctgaagagg	gaaacacagg	gcatgagagg
29521	acggcaggga	gaccggcctt	gtgcggaccc	cctcagctca	gggttctgag	gcctgcagga
29581	gcccggggca	gcgccatcac	ggcggtgact	cctaaatagg	cttcagcaga	tgggggaagg
29641	gcgaagtgga	aagccgcagc	tctctggggt	ttttaccctc	cggtgaaaac	gtaggccgaa
29701	aatcgagct	gcgaagggcc	cgcggctctg	tgcgggtcca	tccccaaagt	tctggcagSa
29761	gccgaatata	tggctgtaga	ggacaacatc	gcacggctgc	gcctgcggat	ccgacacttg
29821	ctgtctcacg	gcgagatggc	tgccttgacc	ggacgttacg	ccacttccgg	cttctcctga
29881	agttcgctcc	cggcctctct	atctcacgct	agtcgttctg	cctggagggc	tgcacggcgg

```

29941 cttgtccctt tggtagttga atcccgccca ttccaaaaag cgctgacagg gatgtaaagg
30001 gttttttttt tttgtttttt gtttttttcc ccctcgaaga aaacattgga attcacccca
30061 atggacaaaa atttaagtct gaccatacaa aaaaattgtc agaactatgg cgcaacggca
30121 actcgaataa cggtggggaa gtttaattgtc ctggctaata aaaaatgtat ataacatttc
30181 ctatccttaa agagctcaca acctcactga taataaaaag taaaaagaaa acaagcagta
30241 taacatatga ttacgccaca atgaactaca gaagggaaaa tcaaggcgtg ctgaagtcce
30301 actaagaaac aactgcggaa agagccatgt gacaacagtg catgaactgg gagtggcaga
30361 actgaatata aatgcatgtg taaacacaag ctgtttgttt tgcttagtgt tccttgatcat
30421 tctacacgct tgaagatcag ctagcgttct tgctgacagg taaggaggac gcgcttactg
30481 agtgccaagc actgctcagg cactgattct gtcaatctct gtcaatctcc cgacagccca
30541 agggtaagca ctgttatcat tattcaattt tacagaaaaa aaatgcgggg gagaggctcag
30601 gtaacttgtc gaaggtaacg ccgctagtgt ctttaaaaca caacaacaac aacaacaaaa
30661 cacactcaca catatacaca cacacgccat ttaaaaaatc atctttccta cgtccagcaa
30721 gggccaatta gagatggctg tggcagggcg gccccgccc ggaactcctc aagagcttcg
30781 cccctcctta cctatggaaa cacaggaagt gacctatgct cacacttctc aYggcctcgg
30841 ccctagtggg agcaactcgc tgaagccgag ggcagaactg gcggaagtga cattatcaac
30901 gcgcgccagg ggttcagtga ggtcgggcag gttcgtgtg gcgggcgcct gggccgcccg
30961 ctgtttaact tcgcttcgcg tggcccatag tgatctttgc agtgaccag gtaacagatt
31021 gtactctttt ctgacggttc gggcgaaggc caccactgca ctgaggcctg ggggcaattg
31081 tggggaagag actaggaatt ggcgcgcgtg caggcccctc gggggacgtt cctcccttcc
31141 gtgctgccgc cgttcggccc tgaacggccc actcggccgc cactcccgcg tgggtccctc
31201 ctctgctgtg tttcgcaggc agcttcccat cgtacgattg tggggctcag ggtactactg
31261 gctggctggg cggcggcagg cgggacagga cagtccttg catcgaagac cctaagttta
31321 ccctgccctg tcctgccatc cgcttctct ccatgttaga agcagattca ccagatctg
31381 tgcccgcctg ttttgcgtgc aacattgaga cttaaattt ttgtcagaag cctgagacag
31441 cgggcacggg agcgcttaag atataatata caccacttta tttgcagggt ctcccgtctc
31501 tcggttcagg ccatcatggt tttccaaatc tctaggtaga cttttctgtg aaaagactgt
31561 gcttcattta gttatacaga cactagaagg ctatgcagaa ttaatttgat tgccctccaa
31621 aaatatcgga tttgatgttt caatttccag gagatgaaga taccagcaa acaactcttt
31681 tctgaggata aattagtga gtaatcactg tgcggtttct tctgtagact tacttgcaa
31741 aagtggccctg aagccaccga agtccttga taaatctcta atcatactta taactgcttt
31801 aaatcctgcc gtcattatct cttgcctcaa ccttagattc ctgaaacgaa acttccgtcc
31861 tccagtttta ctctctcaa attcatctag tcttgccaaa ttagatctgt tcatactgca
31921 cttccgaaat tccataactg ttattattgc ctatgcaata acattgaaaa ctccgtatag
31981 tatgagccca ccaatatgtg ctgtctcatc tgctgcagtg acctctata cagtataact
32041 aagcttggtg cctgcatact gcatgctttt tcaatctgtc tctttctgct tgatttctct
32101 tttgtctgaa gccctgatgt gtaaatctct actcaccttg tgagacccaa gttagatggt
32161 ccctgctttg tgaaaacact gcgctctctt cacagtgatt ggctgttagt ctatattgtc
32221 ttctcttcca ggggtgtata tgggctcatt catgatcaca tactgtattc caggcatagt
32281 gctagtgca gagatcaca agacatgtag gctggtttct gcattcaagg aacttagctt
32341 agaccatacc tgcgtttata atactatggt ttacagttag tatttgcata cccttcatat
32401 tgaacacttt gatgccaagg actatatcct cctatcttta tatcctcatc tgcaggactt
32461 ctgttattgt tattatagga taactgtcaa aaaaaaagta ttttttaaaa aatatctctg
32521 atatatttat ttccagaagc agagcttgct ttcttttttg gtctgttttt cagtgtatgag
32581 tatgtaggat agatagtctt tgggggcctt tgcccttcca aagtgatcgt cagagtcttt
32641 catacattca gcaaatatct gagtgcctgt tctgtaccag cacatgcttg aagtgcatat
32701 gcctgaagga tctttggaca tataatttgt aactttgaga cctctaagtt ctatgtgaga
32761 atatgttggt ataaactcat ttcagatgtg tagtgagtaa agcgtatgat atttaagaaa
32821 agtcagataa caggcacagt ttgcattaat gtgttctaaa gaggttaaggt tattacattt
32881 ataaaaattc agggctttat ctttgcggg cttttttttt ttttaacagt tcattacagt
32941 aggagcttga taaatgatca ctctgaagta tattggattg aatttgatat ttacttaatt
33001 ttttgcccaa gacattgtag aggatgtaa attggaatat ttaaagatct aaactttgcc
33061 taacagtgtc gtgtatacag tgcttagtga atattctgct ctgatattac attttgctta
33121 ggaattattt ttctctaggt gtttttctc aaaaagtttt aatgctggtt atgacagctc
33181 gattttgagc attttcogat tatttaaaaa tgtaacaaaa tgatttttgt tttgtggcg
33241 attttacatg caatcgccgg aaacatggaa ggaataaaac tttaggatta taaggtaaaa
33301 acaaatgtat tccaaaatag cttcattggg tttcatgttt gtgttttgta tagccataga
33361 actggcttat aggactgtac aggttacctg gatccttaa ttaaacttta gacttttttc
33421 caaagcagca tcaactgttc ttggcgtgtg aagataaacc aaggaattga ggaagttgct
33481 gagaagagtg tctgagagt gctctaggaa aaaattgaat agtgagacga gttccagcgc
33541 aagggtttct ggtttgcca gaagaaagt aacatcatgg atcagaacaa cagcctcca
33601 ccttacgctc agggcttggc ctcccctcag gtaatatagc agggaggaga gaataggag
33661 ggcggaaatc tgaactgcaa gagatggtat caaaaggcaa ggaaggcat ttaatgatc
33721 gtttttgaaa atggtttaat atgtttttta agccttattt tgttgagaag ttctattagg

```

33781	ctttgaatag	gcacaatggt	gtttattttg	gaagtctgga	agctaagtta	tatatattatg
33841	aaacacctaa	tcttttgata	aacacttatg	agaatgttcc	aaatgactat	ataggggtctt
33901	tgattttgaa	aatcaccctc	accaaattat	tttctagttt	tatttttccc	cttttacatt
33961	tcaagctctc	aaccacccat	ctactttaaa	aatttttcag	cctgggtgca	gtgggtcacg
34021	cctgtaatcc	cagcactttg	ggaggccaag	gtgggcggat	cacctgaggt	caggagttcg
34081	agaccagcct	ggccaacgtg	gtgaaacctc	atctctacaa	aatatataaaa	aattagccag
34141	gtgtttgtgc	gggtgcctgt	aatcccagct	actcgggagg	ctgaggcaag	gagaattgct
34201	tgaacccggg	aggtggaggt	tgcagtgagc	agaggtcgtg	ccactgcact	ccagcctggg
34261	agacagagcg	agactctgtc	tcaaaaaaaaa	aaaaaaaaaaa	aaaaaaaaatca	ataacttttg
34321	ttgaaatata	catacaaaaa	aactcatttt	aagtgtctgt	tttgataatt	atatataatg
34381	tataaattac	agaacatttc	agttacctta	aaaagtcctt	tcttcccctt	tatagtcact
34441	ctgctggccc	caggtaacta	ctgcatctgc	ttttcaatgc	tgaagattag	ttttgtctat
34501	tctagaatth	catatagatg	gaatcagagt	gtatgctttt	ttgtgtatgt	ctgacttctt
34561	agcccagtg	actgtttgta	tatcagtagt	taatccattg	tatagctaag	tatcactcca
34621	ttgtttggat	gttccacagt	tcattccattc	tccagttgct	cacatttggtg	gtctttccag
34681	tttggagcta	ttgcgaataa	aagcactgta	aacatttggtg	tagactttga	atgcactggt
34741	tttacttctc	atgggtaaat	acttaggagt	aggattgcta	ggtcctatat	tgggtatatgt
34801	ataactttat	aagaaactgc	caaactgttt	tttgaagtgg	ctgtattggt	ttgcagcata
34861	agagatttaa	gttgctccac	atcctcacca	acactttctg	ctgtcagctc	ttttactttc
34921	actctagtga	ttgcttagta	gtatcttatt	gtggttttga	tttttatttg	cctgattact
34981	aatgtttctg	agcaccttgg	caagtctctg	tcagccactc	atacacaggc	ccacttccact
35041	ttactgcact	tcactttact	gcactgtttga	caaattgaag	gttggtggtaa	acctgtaccc
35101	agcaagtctg	ttggcattat	ttttccaaaa	gtgtgtactc	acttcatgtc	ttgggggttag
35161	attttggtaa	tttttgcagg	atttcaaact	ttttcattat	tatccgttat	tgttatcaaa
35221	cttttttttt	tctttttttt	tgtatttttt	ttttaatttt	ttttttcagt	cactctgttg
35281	cccaggctgt	agtgagtggt	tgcagtcttg	gctcactgca	aactccatct	cccggttca
35341	agcgattctc	ctgcctcagc	ctcccaagta	gctgggatta	caggcatgctg	ccaccacacc
35401	cagctaaatt	ttgtattttt	agtggagacg	gagtttcacc	atgctgggtg	caaactactg
35461	gctgcaagt	atccatccgc	cttggcctcc	caaagtgttg	ggattacagg	catgagccac
35521	tgtgcctagt	caaacctttt	cattattatc	tgttattgtg	atcagtgatc	ttttagttaa
35581	ctattgtaat	tattattgaa	tgccataaac	tgccactgta	taatacagga	ctttaatgga
35641	taaatgtctg	cagcgtgacc	aacctattcc	ccatctctct	ccctctgctt	ggtcttccct
35701	attccttgag	acacaacaat	atggaattaa	ggctaattaa	taaccttaca	gtgactttta
35761	attaagtgtt	cagatgaagg	gaagagctgc	acatctctca	tttgaacca	aaagctataa
35821	atgattatac	ttagtgaaga	aggcatgttg	aaagctgaga	caggctaaaa	gctaggcctc
35881	ttgcacaaaa	tagttagcca	agttgtgaat	gcaagggaag	agttctttaa	ggaaattaga
35941	agtgtacttc	cagtaaacc	acaagtgata	agaaagcaaa	acagccttat	tgcttatatg
36001	gagaaagttt	gaatggctctg	gatagaagat	cataccacc	acacatcccc	ttaagccaaa
36061	gtgtaataca	gagcaaggcc	ttaactctct	tcaattctgt	aaaggctgag	agcggagagg
36121	agctgttaga	agaaagtttt	gaaactagca	gaggtttatt	cagagcttta	aggagagagg
36181	ccttctccac	aacataaaaa	tgcaacatga	agcagcaagt	tatctggaag	atctagctga
36241	gataattgat	gaaggtgact	ataactaaata	atagattttt	aatgtagatg	aaacagcctt
36301	ctactggaag	aagatgctat	ctaggacttt	tatagctaga	gaggagaaat	cagtgcctag
36361	cctcaaagct	tcagaggact	ggctgactag	ttaggggcta	atagagctgg	tgagtttaaa
36421	ttgaagccaa	tgctcatggt	ctgctctaaa	accatagagc	cttaagaatt	acgttaaate
36481	tactctgcct	gtcctcagta	aacagaacaa	caaaaacctga	tgagagcacg	tctgtttaca
36541	gcatgattta	ctggatattt	taagctcttt	gagatctgct	cagaaaaaaa	gtttaatttc
36601	aaaatattac	tcactgacag	tgtaactagt	tgtccacaag	ctctgatgga	gaagaacaag
36661	gagattaata	ttgttttcat	gcctgcttaa	ataatataatc	cattcttcag	cccatggatc
36721	aaggagtaat	ttcaactttc	atttcttact	atttaagaaa	tacagcaggg	catggtggct
36781	catgcttgta	atctcagtac	tttgggaggc	caagggtgga	ggatcactca	aagctaggag
36841	ctcaagacca	acctgggtaa	caaaacaagt	ccctgttgct	acaaaaaaaa	attttttttt
36901	aattagctgg	tcattggtggc	atgtgcctgt	agtcccagct	acttgggagg	ctgaggcggg
36961	agggtcactt	gagtccagaa	gttaaaggct	acaatggaga	ccctgtctca	aagaaaggaa
37021	gcagggaggg	acatatgctg	tagctgccat	agatagtgat	tccctgatg	aacttgggca
37081	caggtaattg	aaaaccttct	gaaaaggatt	caccattcta	gatgccatta	agaatgttca
37141	tgattcatgg	gaggagggtg	aaatatcaac	atgaataaga	gtttggaaga	ggttgattcc
37201	aacctcgtg	gatgactttg	agaggttcta	gacttcagtg	ctggaagtta	ctgcaggtgt
37261	agtggaaata	gcaagtgaac	tagaattaga	agtgaacctg	aagatgtgac	tgaattgctg
37321	caattttttg	ataaaacctg	aacagatgag	gagttgcttc	ttgtgagtaa	gcaaaagaa
37381	tggtttcttg	agatagaatt	gactcctggg	gaagaactga	tgacttttaga	atattacata
37441	aacttagttg	ataaagcagc	agcagggttt	gagaggattg	actccaattt	tgaataaagt
37501	tctagtgtgg	gtaaaatgct	gtcaaatagt	atcatatgct	acagagacat	cttcagtgaa
37561	aggaagagtc	agtcagtggt	gcaaaacttcg	tcagtccttat	tttaagaaat	tgccacagct

37621	accaccctga	tcagtcagca	gccatcaaca	tcgaggcaag	atcctctgtc	agcaaaaaga
37681	ttatgatttg	ctgcaggctc	acatgattgt	tagcattttt	agcaataaag	cattttttaa
37741	taaagttata	tacataattat	tagacataat	gctattgcac	acttaataga	cttttagtgct
37801	aacataaact	ttgtaggcac	tgggaaacca	aaaaattgat	gccgcttgct	ttattgagat
37861	ggtctggaac	ctaacctgta	gtatctccga	ggtatgcctg	tatcttcatt	tgtaatatgt
37921	ccttcacatc	ttttgccctt	ttttattatt	ttatttggtg	atcttctttt	atggagtgtg
37981	cagagctctt	tattattctg	tttaccagtc	ctttctcaga	tgtatgtatt	atagttattt
38041	ttttcccagt	ctggcctlgcc	ttttaatttt	ctcaatgggtg	tctttcaaag	aacagaagtt
38101	tttaattttt	ccgaagttca	gtttatccat	ttttcttcat	gtttatccac	tgtgtgggtat
38161	taaagaagaa	agcaatgtgt	ataagaatag	ctggttcttc	cgtaattaat	gtttaataac
38221	cccattattc	tccgaaggca	tctgtctttg	cacacctgac	ctgctgttcc	accaagaaag
38281	ttccacaaac	acttagcagc	agccagccta	acctgttttt	ctccttgctt	tccacaggg
38341	gccatgactc	ccggaatccc	tatctttagt	ccaatgatgc	cttatggcac	tggactgacc
38401	ccacagccta	ttcagaacac	caatagtctg	tctatttttg	aagagcaaca	aaggcagcag
38461	cagcaacaac	aacagcagca	gcagcagcag	cagcagcaac	agcaacagca	gcagcagcag
38521	cagcagcagc	agcagcagc	gcagcagcag	cagcagcagc	agcaacaggg	agtggcagct
38581	gcagccgttc	agcagtcaac	gtcccagcag	gcaacacagg	gaacctcagg	ccaggcacca
38641	cagctcttcc	actcacagac	tctcacaa	gcacccttgc	cgggcaccac	tccactgtat
38701	ccctccccca	tgactcccat	gacccccatc	actcctgcca	cgccagcttc	ggagagttct
38761	gggtattgtac	cgcagctgca	gtgagtactt	cgtgttttat	gtttctctcc	acttaggagt
38821	cccttttagt	talgttctctg	ctctgttttc	agatggatcc	ttttattaag	ggaggagagt
38881	gcactaacgg	taattgtgta	tcaaaatttg	ctttatctca	catttgggaa	agggaaagca
38941	agctatctta	gtcagtgtcc	tcagtaaaag	gctcttaaca	ggtttagaaa	tgtggtcatt
39001	tgtgtttaca	tacctgagcc	aataaaattt	aatctgactt	tcactgtcgt	tattattata
39061	ttatagacat	ttccctgtat	ctgatatacg	taaatcacia	tgttaggtag	tctctttccc
39121	ttatgctatt	ttaggctctt	agtcacataa	tcagtataat	ttctggtagt	tctgtttttt
39181	tgtttttact	atgggtgctg	catataaaatc	tttgaaggtc	tgtgtccttc	ccacaaaatg
39241	aagacgactg	tttttgtcat	aaatggattt	ttctacctaa	atgaagtggg	ttctatatgt
39301	aacagtgtag	taggggtagg	aaataatttc	atcttctctga	aaaaccagca	aatattttcca
39361	ataaattcca	gtcaataaga	agtgtatagc	ttttcatttt	agaaaagctta	tgaccaaaat
39421	aaaaggttac	ttgcagctctc	tgcatctcct	cagttttctt	gtacagatac	cttctctctt
39481	catactccct	ttagatctag	tatttccctat	ttgcatttat	tcaccttggt	acataatagg
39541	ttgttggtta	taagctatct	ctcatcattt	ttggaactaa	gtgggtgtga	atcgtggatg
39601	ctgactttca	cccaaagacc	taaatcttgc	ctccacaatt	ttgtttgacc	ccacacacaa
39661	ttttaaaaaa	aaagtaattt	tagtgatttt	agagagagcc	cttttttagct	cacctttgtc
39721	tataccactc	ccatgtgtcc	agcatgcttg	tctgactccc	aggaatattt	aagtttgcat
39781	ttctgattta	aataataaat	tataaacagt	acagtcaggt	gtttgtttct	tgcgagtggc
39841	tttgctgttc	tgatactccc	cagagcatcc	agtcacccac	attcctgaat	ttccactatc
39901	tactaagaac	cctcccgggt	ggtagtccag	tgccctctcc	atcttgtagg	cttaagaggg
39961	ataatggtaa	atatttaatt	aatgaaaata	tataagaata	catttaacaa	ttatgtgact
40021	aaagttgtga	tagatatact	aaaatctact	tggtatagtt	cataagtgga	caaagtctaa
40081	aggcaggtgt	gtatacaagg	gtgattggaa	ggaaatagat	gatattggaaa	tgaaaacatt
40141	cattgtgtgg	atcaacctaa	ttggctttac	agttgtgtat	catagactgt	acaagaaaaa
40201	gaaccagaag	caacaatgag	ctgaaacctc	ttgtatttga	acctatgtat	ttgaaatctg
40261	gactcgatat	tcttgggatg	gtgattttcc	ggtccatgcg	agacagttgt	tgagcagtta
40321	attgcattca	aagggcattg	cccatttcaa	ctatatatac	tttcataagc	agaaaaatgt
40381	ataatgaaaa	tgtgggtttg	ttgtttaaat	ttttattaaa	atcttctaata	gagttgcttt
40441	ttagtatatc	tttatccttg	tataagttat	tcgttaagtg	atttaagaac	actgagaaat
40501	gaaaagggtc	agtggaccca	gatggctcca	tttatcctta	tttactgaga	gttaatttaga
40561	agagaaaagca	gggttttgcc	tttttttttt	gctaaagaca	cttagctctt	gtttttcaaa
40621	agatacttac	tttgagagaa	ctgggcagtt	catgctgctt	tgttttcatt	ataaatagtt
40681	acttaggaat	agggtggtgg	tggtagcttt	tcagagctca	cagatgagaa	ttgagtttgt
40741	ctgtgagaag	gctaagggca	ggctctggag	acagacagac	tcaaatcctg	gctcctccac
40801	ttgggagcca	ttgcgtcctg	agcaagtgtg	actctacctt	taagcctcag	cttattttatc
40861	caaatgtggg	aaactgggta	actctacctt	cagtgcgtta	tatgtattaa	atggaataat
40921	ctatataaaa	tgctcaacat	gtgcatgata	cttggttaaac	aataaatgtg	agcaatttaa
40981	aaaaagtaaa	gcctggttga	aataatcaga	tgtctgcata	atcttctaacg	cctcatccaa
41041	tgaaacttaa	gtaattttaa	tagtcgtgtt	ttctttttta	atctcttaca	gaaatattgt
41101	atccacagtg	aatcttggtt	gtaaaactga	cctaaagacc	attgcacttc	gtgcccgaag
41161	gcgccgaatg	aatcccaagg	ttagatctat	tttaagtgtg	ttcttttttt	ttcttttttt
41221	gtctcctctg	ccgtgtctct	atattttaat	gtatttcacg	ctataaaaaca	aatgtctgta
41281	gatcaggcca	ggcacagtgg	ctaatacctg	taatcccagc	actttggggag	gctgaggaag
41341	gagaattttct	tgaacctagg	agttcaccac	cagcccaggc	aacatggcaa	gaccctgtct
41401	cttaaaaaaa	aaaaaaaaat	tggccgggct	tgggtggtca	cacgtgtaat	cccagcactt

41461	tgggaggcgg	agggtgggagg	atcatctgag	gtcaggaggtt	caagaccatc	ctggccaaca
41521	tagtgaaacc	ccgtctctac	taaaaatata	aaaatttgcc	gggagtgggtg	gcatgcgccct
41581	gtaatcccg	ctacttggga	ggctgaggca	ggagaatcac	ttgatcctgg	gaggcagagg
41641	ttgcagtgag	cgggagattgc	gccattgcat	tccagcctgg	caacagagcg	agactccgcc
41701	tcaaaaaaaaa	aatctgtata	tcaaagagtt	tgtgttatgc	ttattccttg	acaccaataa
41761	aatgaagatc	taaagtaaaa	tgtgcattgt	tttgtatcct	ttattgagtg	tctgtgatata
41821	agaatccct	gggagtacag	cagtctataa	aataagctaa	gttatatatg	tattgtattg
41881	ggatcatgtg	aagaattttg	tatatgtata	gggatgatgt	tagccagcta	tagtaggcaa
41941	agtaattttg	ataaacatta	gctgggttga	gagaaatgtt	aagtaatgga	cttaggggtgc
42001	atttttaaat	ccctgaaaca	taatgcctga	gatctgaaag	acagctaagg	gtctgaaaaat
42061	caccttttaa	ccactcataa	tgagtctgta	tgacatgtag	cattctaaga	attaatcttt
42121	catctattat	aatcacattt	agttgaatac	ataattttat	tactttacgt	tcttaagtga
42181	tatttaacca	ataaaaatag	tagaggaaat	tctggaaatt	caaaatagct	aggcttttcc
42241	tgggacgtat	ataaactact	atatactgct	agctttttta	gatgttgata	aaatcacagg
42301	gaagacatag	agcagggtttt	aaactaaaat	tgttaaaagc	caggcatggt	gactcagacc
42361	tgtaattcca	gcactttggg	aggccaaggt	gggaggatca	cttgaggcca	acagttcaag
42421	acgagcctgg	gcaacatagc	aaggccttgt	ttcatcagaa	aatttaaaaa	attatctggg
42481	gttggtggca	cacacctata	atcccagcta	cttgggaggc	caagacagaa	ggatctcttg
42541	agctcaggaa	ttcgaggctg	cagtgcagcca	tgattgcacc	actgcactcc	agtcctgggca
42601	acaaagttag	actcttgtct	ctaaaaaaa	aaattaataa	aataaaaaaa	taaaattgtt
42661	aaagcaatct	agtttgtcaa	agagtcgtct	taattagaaa	catgtaggtt	tttattgtat
42721	aattaaatgt	attaaaatat	ttaagttgtt	tcaaagtaat	tatttgcttt	ttatcatgta
42781	cctaataact	gaggccatta	attcaaccca	cagttgaaat	attttcccct	tgagttacca
42841	ccttagcaag	aaaattttct	gaaagaatta	acatgtttca	agcataaaaa	gcagaaaata
42901	aaagttaaaa	ctgtgttttt	ttcctaattt	gcatttatgt	gtttaatata	ttaatattgt
42961	ttctaagagc	caattgaaag	catgatttta	cttcattacg	tattgatagg	cttgtaactc
43021	gatttattaa	ttccattaag	gtagaaaagg	tttattgggt	tatttgcctc	aaaattcata
43081	ttaataaaac	aaaatgtatc	agagaagcta	tagtcatgat	aagcagcaaa	tataaagtag
43141	aaacagaagc	acttgtattt	attgagaaca	taaccaaat	ttgcaaacat	ttctcttgat
43201	ggaatgtcta	tgtgaaagaa	catacagatt	tggtataact	tagtcacaat	tgattatttt
43261	gtgctcctta	tgcaaaggaa	tatttggttac	atgtgggtga	tgcaaatcct	ttacagatat
43321	acagaaatac	agaacaaata	ttttgataac	tagatgtact	atgtccttcc	taccagttgt
43381	gatttttttg	tcaatgggtg	gttcgcctaa	caacattgag	cagtttgcca	tgacctcact
43441	aatgattctc	tctgaccatt	gtagcgggtt	gctgcggtaa	tcatgaggat	aagagagcca
43501	ogaaccacgg	cactgatttt	cagttctggg	aaaatggtgt	gcacaggagc	gaagaggtag
43561	ccgtaagaaa	ttcattcttc	tggtctatgg	gttatgaatg	aaaagggtgat	atctcattgt
43621	ttttagggtta	ttagggttagc	actttaacat	gttattattg	ctttcttata	aaaaccattt
43681	taatattgatt	ctattaatta	tttttattta	tttatttatt	atttgattct	attaatgtta
43741	gatagcagtg	atttcatttt	cttaaaaaat	agatatgggc	aggcgtgggtg	tctcaccaag
43801	gcgggcaaat	cacttgaggc	caggagttcg	agaccagcct	gaccaacatg	gtgaaacctc
43861	atctctacta	aaaatacaaa	aattagccag	gcgtggtagt	ggatgtctgt	agtcacagct
43921	acttggggagg	ctgaagcacg	agaatcgctt	aaatccggga	gggtggaggtt	gcagtgcagct
43981	gagatcatgc	tactgtactc	cagcatggga	gacagaacaa	gacactttct	ccagaaaaaa
44041	gaaaaaaatt	aggtgaagatg	taatacaaca	tcagaactat	taataaaaat	ccctttatgt
44101	gaaaaaatgtt	gtaaaataaca	caaatcacatc	ttgaaatgga	agaaaaaatt	gatcagaagt
44161	taaaacactt	ctaagctaata	ggtgtaaggg	ctttgcagac	cttatggggc	ctggcagcgt
44221	gacagcatca	catgtgtttg	ctgcagaaac	ctgccttctc	actggaacta	gctttcatta
44281	atagtgtttg	agttgtgcag	atcttaagga	atgcaccctt	gttataaaaat	gtcattgcct
44341	aatattctcca	tctgaagtga	tcctgtctga	aggtttagtac	tctgcggccc	tgactgttct
44401	aaacagagct	tatatacttg	gtagaagtca	aaaactaggg	ggaaaaattag	tgacattagc
44461	ttcataaatg	aactgaacaa	ataaactaaa	actttttaaa	tgaaagagta	tKaaatgtac
44521	ttggagggtat	caacacatag	aagggggctg	tggtatattt	tgcttattct	gaagtttgaa
44581	attgtctttc	agtgatggga	atgccattgg	tgtggacatt	ctggctcctg	atctccgagg
44641	tttttcagaa	tgaggagcaa	ggatgtgttc	ttgcctttca	cttccccttg	gccacaacat
44701	gcagtagtaa	cctctttaat	taagagcggt	tttgtttgct	tgttttcccc	tgcatggaaac
44761	atcagaaact	tttgggtttat	caaggcaagc	ttttcatgca	gcatttagcc	ttttgtccc
44821	agagcatctg	aaaactgaat	attgtatatS	tagttggatg	actatttcat	caagacaaga
44881	aaccaccaat	acatttacct	gaaatttaaa	agcgtagcat	atatatatat	gcttgatttt
44941	tgttcttctc	gatccccctgc	acctgatagt	ctctttatgg	taatattttc	acattctttt
45001	ttgtgttttg	tggtgttttt	gagatggagt	ctcgcctctgc	gcccaggctg	gagtcagtg
45061	gcatgatctc	ggctcactgc	aagctccgcc	tcctgggttc	acgccattct	cttgcctcag
45121	cctcccagag	agcagagact	acaggtgccc	accaccatgc	ctggctaatt	tttgtatttt
45181	tagtagagat	gggggtttcac	tgtgttagcc	aggatggtct	caatctcctg	acctcatgat
45241	ccgcctgcct	cagcctccca	aagtgttggg	attacagggtg	tgagccaccg	tgcccagcct

45301	gcctgcctgc	ctttttctct	tttttttttt	ctttttttct	ctttttttct	tttttctttt
45361	tctttttttt	tttttttttt	tttttttttt	gtggagtcac	ctaggctgga	gtgcagtggc
45421	acaatcttgg	ctcactgcaa	cctccacctc	ccaggttcaa	gtgattctcc	tgccctagcc
45481	tcccaagtag	ctgggattgc	agggtgcctg	caccatgccc	agctaatttt	tgtattttta
45541	gtagagatag	ggtttcacca	tggtgaccaa	gctggctctg	aactcttgac	ctcaggtgag
45601	ctatccgcct	tggectccca	aagtgcctgg	gattagaggc	gtgagccacc	atgcctggcc
45661	cacattctgc	ttttcttatg	taagctctga	actgctaagt	cgtagtttat	tcaacaaatg
45721	acataggaat	gtctattcat	gatgagtctt	ggataaaaag	aggatagaat	tagtgaatac
45781	attgttcaat	aataaatctc	atcaacattt	tctgatcaaa	atgaagtttg	ttagttttcc
45841	tctcagtaga	aatgcattgg	ctaaaataca	gaaatagtga	tgaYtgatga	tggtgataat
45901	tcacatccat	aaaatctagg	gctacaataa	tttggcggat	tgaaaggcca	ttttggcagg
45961	cctacagttt	tctgtcaagg	atccaggaat	actttataag	gaattgtgaa	tgccctgtcag
46021	tcttttctcc	tattgcaaga	aggctgacca	gtttacactt	tattagttta	ctgttttgga
46081	ctttttataa	gttatttagt	tataataagta	ttttagctgg	ctctgagtat	gaataactca
46141	cttttttctc	ttccctagtg	aagaacagtc	cagactggca	gcaagaaaaa	atgctagagt
46201	tgtacagaag	ttgggttttc	cagctaagtt	cttgacttcc	aagattcaga	aRatggtggg
46261	gagctgtgat	gtgaagtttc	ctataagggt	agaaggcctt	gtgctcacc	accaacaatt
46321	tagtaggtaa	gtctgaaatg	tattRtgatt	gttattggca	acagttcatt	tataatctaa
46381	acattgttca	gaataaaaac	catgcaaaa	attcagtata	tgagaacagt	tgacatgggt
46441	atagttgtat	gtattcttgc	attgtcttcc	tgatgttctc	agtcataatt	atcacctcca
46501	ccagctctcg	cttcccttat	cactttgcgg	taccataaac	tcccttttac	tggaatgaat
46561	ttgattctac	ttctgtatgt	tttatcgttt	tattgctgaa	tacacttgga	atgcatgaat
46621	tgaccctaac	cttgtatcaa	ttttttttcc	cagtttggtg	cttctctttt	agtcaacatt
46681	gtgtcagatc	taccagcaaa	gtgtgaagtt	gagcgatagg	aacaactttc	taattatctt
46741	ccctgtact	tgccaggtgaa	gactcacagg	caggcagccc	tgcccacctc	actgcttcat
46801	ctcaggtctc	ttcgggtgct	gttcacaggc	acattagtct	tggtgtgccc	tggtgtctac
46861	ttgggttctc	ctgcccctag	cctttgcact	tgccactttt	atgcctagta	tgctcatttc
46921	ccttttcacc	taggaagctc	ccacttgagt	tccatatgtc	taccatttcc	actcttcctc
46981	agagaagcct	gtcctgaatt	ttgtcttagg	tcagacacac	agaaacagag	gcattcctgt
47041	ttcagctctc	ccgaactgag	aaaaatgaga	ttactaggct	tagaagaata	tgaccaggc
47101	ctggcctggg	gtgtcacacc	tgtaatccca	gcattttggg	aggccgagcg	agtcagatca
47161	gttgagttca	ggagttcaag	accagcctgg	ccaacatggg	gaaacctcgt	ctttactaaa
47221	aatacaaaaa	ttagctggac	atggtggtgc	atgcctgtaa	tcccagttac	atgggagggt
47281	gaggcaggac	aatcattcga	accaggaga	tgagggttgc	agtgaagcaa	gatcgtgcca
47341	ctgtactcca	gcctgggcga	cacagcaaga	ctccgtctca	aaaaaaaaaa	gaatatatac
47401	caatagtcca	ctcagtcaga	cagcttaatc	aggtataggg	taatttctcag	gctagtatat
47461	aagtttgatt	aaatttcctg	accacaattg	tcagctagag	aatatttcaa	tttaaggagg
47521	taagatatga	ttaaaagtta	aactgtcagt	attggatctt	agaagtaa	gattattagg
47581	actgtaatat	taattattag	gactgtaaaa	gtaaaaggat	attatctgca	ttagatatca
47641	ttatatctaa	tgatatagag	actgcagaca	taactacagg	gctctttttc	ttaaatcaga
47701	aaatccagat	tcaatagaaa	tagggtaaa	tgataggagg	acaaatagcc	ttccatccag
47761	tggttatcaa	ctgacgacta	caagtcggcc	tcacttgctt	taattattct	attctatcct
47821	ttgatgctgc	ttgaagaact	gtgtttttacc	tcttgactag	tttgtttatt	cagtattttt
47881	ccttgtacag	gtcctcattt	tatctaaaa	cacacaaagc	tcttgatttc	taaacttttt
47941	gcaatttttc	ttctagtatt	gagccagagt	tatttctctg	tttaattctac	agaatgatca
48001	aaccaggaat	tggtctcctt	atttttggtt	ctggaaaagt	tgattttaaca	ggtaagttgt
48061	aacaggaagt	agtatctgaa	agtttgtaag	tgttttgagt	atggcatttt	ctcagtgtctg
48121	aaaagaaatt	tcagtgttcg	gacagtgggc	tagcttcttg	tacaaaggcc	tcccacccaa
48181	agtctgatga	gaaacgtgcc	cactaaaggc	acagtgaag	cagggaagtc	tgaccacagc
48241	tctgcaagca	gacttccatt	tacagtgaag	aggtagcat	tgcatgaa	aaaagatggc
48301	gttttcactt	ggaattagtt	atctgaagct	ttaggattcc	tcagcaatat	gattatgaga
48361	caagaaagga	agattcagaa	atgagtctag	ttgaaggcag	caattcagag	aagaagattc
48421	agttgttatc	attgccgtcc	tgcttggttt	atggcctggg	tcaggaccac	ggagagaagt
48481	tggaatacat	gcctcttgag	ctatagaatg	agacgtgga	gtcactaaga	tgatttttta
48541	aaagtattgt	ttataaaca	aaaataaagt	tgtgacaagg	gattccacYa	ttaatgtttt
48601	catgcctgtg	ccttaatctg	actgggtatg	gtgagaattg	tgcttgacgc	tttaaggtaa
48661	gaatttttacc	atcttaatat	gttaagaagt	gccatttcag	tctctcatct	ctactccaac
48721	ttgtcttctt	aggtgctaaa	gtcagagcag	aaatttatga	agcatttgaa	aacatctacc
48781	ctattctaaa	gggattcagg	aagacgcagt	aatggctctc	atgtaccctt	gcctccccc
48841	cccccttctt	tttttttttt	taacacaaat	agtttggttt	gggtacctta	aatgggtgtg
48901	ttgtgagaag	atggatgttg	agttgcaggg	tggtggacca	ggtgatgccc	ttctgtaagt
48961	gcccaccgcg	ggatgccggg	aaggggcatt	atttgtgcac	tgagaacacc	gcgcagcgtg
49021	actgtgagtt	gctcataccg	tgctgctatc	tgggcagcgc	tgcccattta	tttatatgta
49081	gatttttaac	actgctgttg	acaagttggt	ttgagggaga	aaactttaag	tgttaaagcc

49141	acctctataa	ttgattggac	tttttaattt	taatgttttt	cccatgaac	cacagttttt
49201	atattttctac	cagaaaagta	aaaatctttt	ttaaaagtgt	tgtttttcta	atttataact
49261	cctaggggtt	atttctgtgc	cagacacatt	ccacctctcc	agtattgcag	gacagaatat
49321	atgtgttaat	gaaaaatgaat	ggctgtacat	atttttttct	ttcttcagag	tactctgtac
49381	aataaatgca	gtttataaaa	gtgttagatt	gttggtatac	cttgtaagag	tcatgtgatc
49441	atactgtttt	ctacaaagtt	gtatttttaga	tataatgcct	gaaaccattt	tggtgtttgc
49501	ttcagtcagt	atttcattgt	atgctgcaat	gaaaacagat	taatgatctg	agaaccttcc
49561	atatattgag	caactcctgt	tttctaggta	ttttgcatac	aatgcctgga	atcctcacaa
49621	agcttcagtt	acgtttttgtt	cctctgttgg	aggtagaggag	ataaggaagc	cccagctgag
49681	ggacttggct	gaggttacac	agctagtaag	tggtaaaaat	gagtgagtcc	ttcaggtgta
49741	gaagctgggtg	ccctatccac	aggctgccaa	ctctctgcag	taactttttt	ttgtttgttt
49801	tgcagttttt	ttctcatgga	ctatcagggtg	gacatttgtg	ggttcttagg	ttttattgtt
49861	agagtgggtt	gttggttttta	attgtaaaag	tacatcttca	gctgactcag	gaataaaatc
49921	agaaagggga	ggtccttcct	tcctttctct	tctgcctctt	tccccaagat	aacgaccaga
49981	ttagtgggta	tcttctaagc	cctgttctgt	ctatcttctg	ggcatggcat	ttgatcccat
50041	ttttggaaaa	aagaaataat	acattcaatt	ttatagcttt	cccttttttc	agtcaatata
50101	tcataaacct	atttccatga	ttgggaatct	aggatactct	tcttaggtac	cttcacttac
50161	tgagctattc	ctgtaattat	ggtcatttac	atgacctcta	gttttttcaat	atcatcaaca
50221	gcgctgcag	gaacatcctt	gtatgtgtat	ttttcggatg	agtgcaggta	ttcatgtagt
50281	aaactttcct	agaagaggaa	ctaaagacta	tacattctga	gttttaatca	gttgatacca
50341	gattgcagtc	caaagagaac	ggtgcagtga	atgccctttt	cttgcaattt	ccaatcatta
50401	gtgatgccaa	taatgtatgg	attaatttga	aagacatcat	tccaatatta	tctcctatct
50461	aagaaagaag	tatctcttca	tgtatgtagg	tcttatgttc	ctcagtaaa	ttttgtactt
50521	tttttcacac	agggctttta	tatttcataa	gtattacttc	cttaattcct	aagtatagta
50581	tcatttctgt	tttattactg	gaagtttttc	ctaactttt	aaaagttatt	gttgataata
50641	tttttcttat	tccgagttta	ctgaactcct	attactgctg	atttagttga	ttgcttgtaa
50701	agttctaggc	atgtaatcat	ctgtaaatag	taattttatg	ttttcctttc	caacatttat
50761	ttttcctatc	agtgcctagg	actttaatgt	atataactta	aaaaaaacat	ctcttagagt
50821	tgtagctaca	tatacaggaa	atctaacaaa	tgtgtagcat	aatgtattat	acaaaggcag
50881	acacctttgc	agccaccaac	aaggtcaaga	aacaattttg	ctgcctgtcc	tagaagcccc
50941	tcttctatgt	ccctatccaga	cacacacttc	tggcttccct	caagcagtg	ctattatcct
51001	gactctcacg	catttaaaga	taattgaagt	tctctgtcca	tctcttttct	ttccaattta
51061	caggccattt	agtgtgtaga	gcttctcata	gtctggggtt	tgctagtgtg	aaattcatgg
51121	tgtagtttcc	ctattctctg	tatttctctg	aaattggaag	ctgctgtgta	attcctagat
51181	ctattaatte	attagtgggt	tgcaaaatga	catttttagtc	atttcatttc	tttttcatte
51241	attaatttga	atacttttat	aaagtgttgt	gctccatata	gcagaatact	ttccccttt
51301	aagttttcaa	gataaaatga	gttcatatta	atatgtccaa	ttcaaagctc	atagggttat
51361	tttaccaatt	gtatatccaa	ttcaaattca	aagcttatag	ggtttttatt	taaccaattc
51421	tgtattacac	tcttccctct	acactgagaa	tttttaattc	ttaaagacat	aggggatgag
51481	gaattagaat	gtcccataat	tactcattta	ctttacgtat	ttactttatc	attactttat
51541	ctgttatatg	tgcaattttt	tttgagacac	ggtctctgtt	gccaggctg	gagtgacgtg
51601	gcacaatcat	ggtttgtctg	agtctcaaac	tctgtgggtc	aagtgatcca	cctcagcttc
51661	tggagtagct	agccaggact	acgggtgcac	accaacacac	ccagctaatt	ttttttaatt
51721	tttgtagaga	caaagtcact	caactgttgc	ccaggctgtt	cttgaactct	taggctcaag
51781	ccatccctcc	acctccacca	taacattctt	aagagtaaca	aaaacactat	caccaatatg
51841	attgccaaaa	acacttgga	ctttttttt	ttttgctgtg	atatctgaaa	ttgccaaagg
51901	atattcagta	tctgaaataa	aaaggcaaa	ctgaatatgc	tgctctctac	agcagagggga
51961	gctgtctgtg	ctggacagta	tctgaaccaa	gcagatctta	aaactttgta	ggtgttgaga
52021	aatggtggat	gcatggactg	gcaccgtctg	tggagccatg	attatgtagg	tgagacttgc
52081	tcatttatctt	gtagtgtttt	aaaatgtctt	cacatttcta	aaggcaactt	gcttaatgca
52141	ttttttaatt	taaatttttt	atgttgcata	gtttatttta	aatatagttg	ctatttttta
52201	acacagatgc	caagtcgggtg	ctgtgagatt	ttctttctgg	tgatttggac	cagtttgtct
52261	ccctcttgat	atatccatcc	caaattggaa	ggccctgtaa	actgttacga	tcatctccag
52321	aggttaactg	gaatatacac	caatgacagc	ttgcctgggt	atgccaaaat	acctgcaaga
52381	atgtccacat	catctggtga	tgtcccaaaa	taacagtttt	taccatagaa	agatcgttaa
52441	catgtttgct	ttaaaagtca	ttagcagtc	taacgtactt	acagattctg	ccataaagg
52501	ataatacata	atttttagata	taaaggcccc	actagtccag	gtttccttat	gccactgtgc
52561	ttcctactaa	gtgttgcgac	cagctcttgt	cactagtgtg	tgacaactta	ctccagtagc
52621	cacagggctg	tgacaccata	gttatagggtg	attttcatag	atttagccat	cccaggttcg
52681	aaactagtat	cctctagctc	ttaagtagct	gataacctcc	ccatgggaga	aactccatac
52741	tgcagtttcc	catatggtgc	tatgtataac	tatcttatac	aattaatata	aatgtcatat
52801	gtatacttat	ataatatgga	ctactaggac	agaacttttt	aaattacaaa	taaaatagcc
52861	aagtagacaa	ttacattagc	aagtgatgtt	acctatgaaa	acgtgaggat	ttatggtggc
52921	aatgcatttc	agttaacagg	gatgtgttag	gggacaatgt	gagccaatgt	agattaaagg

52981	aataagcctg	agaaatttat	cagaatttagc	cgtcagtatt	caagcactga	tcaacagcaa
53041	tgtgtcttaa	gggcaggcat	cactgggtgct	gagagaactg	ggaattgtca	actgtgagct
53101	gctaggggat	ggaagaaacc	ttagtgtagt	cttaggagcc	gcttgcttaa	acagatgtat
53161	cagaaacata	ataggccaag	ggtcagccct	ttgaaaactg	acttcagggc	cttcctttcc
53221	tcaggctgct	gcctcctagg	ccagaccctt	atthttggctt	acattccata	acccttgtat
53281	gtgcgatagg	gaacctgtat	acaatgttga	cacggaaagg	gaagaccatc	gccttttgcc
53341	tttcagtgtc	tcactctgtaa	gcagggccgc	Yggctgacca	agatcagttc	tgaaggcca
53401	gcctctttaa	attccagttc	tgtgatcaca	aagccactgt	tgcttcccat	cctgccaaact
53461	gtgatactgc	tgcttcagaa	ttactgggtt	tcctgttcat	catactcacc	aacctgaggt
53521	ttgggtatttc	tcaaataattc	tgggcttcca	gtacatacta	gagcctgtga	taatcagcta
53581	atgatcagac	aagtttgggt	ggagttttac	ctaagtattt	ttgtgtttta	aaaacctagg
53641	gtgggaaatg	ctcagagtga	gatgggttga	cttcattagg	catataaccc	atthtttatta
53701	taaaaagaaa	tgacacatata	agtaaaaaga	ccattttagt	tagtcccacc	attcgttgggt
53761	aaccagttca	ctaagtgtac	atctttgcag	atttctgtgc	atacatataa	atathttttac
53821	agaatagga	tcataccatg	aataccacct	ataatgggtat	gtctgtttgtc	aatataggtg
53881	ttataatttaa	gactgtgtag	ccttcctctg	Kggatgtacc	aaaattttatt	taattccctg
53941	tcactggaca	cttttgggttc	actaataagt	agacactgtg	taagcaatcY	gtcaacatct
54001	ctgcacctct	atthtttggta	taagtatttc	cttaggataa	aatcccagaa	atggaattgc
54061	aagggtataaa	agattatttaa	catttttcaa	ggcttttaaga	tgccctttacg	gtgtatctgt
54121	tacatcctgc	ttccacacaa	attcttctgt	atagccagta	cccaagctgc	agctctcage
54181	acaggtgaag	acagccagga	tgcccagtc	aatgctcttg	cccagctctg	cagccttcag
54241	gtagtttagg	agctgaggca	tgacctgaga	agagggtgac	acacagttag	aaagctgctt
54301	cacagcaggg	agcacgagac	cttctcagcc	aggatgatta	tagggatctt	ggtctttcaa
54361	tcctcatact	acaaagcagg	attatagaca	ttatacaatt	aacatgttta	acaatctaaa
54421	acttctttat	gacttcaaag	cccctctcac	cttctgtttg	gtctttttcca	tttgagaaag
54481	aagttcacaa	gtggctgtta	atgaattatt	ttcattacta	atatgccact	caaaagggct
54541	gaggcttcta	tttgggcaac	ttttactttg	tatcattgca	gatgtttgta	ctcttgactc
54601	aagaaacact	aattactagt	aatgaataca	gaaaggacat	ctatcaatgt	agttatagag
54661	accagagagg	aatcttagaa	gtagtctaac	tcaaagagtg	aataggcaga	atagccacct
54721	gatatggaat	cactttatatac	aaatcctgtc	acctcaattt	ggacattgag	agctttggca
54781	taagaacca	agcagagttt	tgtgtattgt	cctcataatt	ccttttttac	ccaaagaaac
54841	aaaccaatat	tagctatgac	tttggttaagg	ttagtgaatc	catagctcaa	gagcatttcc
54901	acctacccta	aatggatttt	gatgctaaca	aatccttttg	ggcagggaa	gacatttatc
54961	tttaatgctt	atatccattt	tttctaacaa	atccacaaac	caagatttaa	cagtaaagac
55021	tcctctcata	aagtatataga	tcaaagactt	taattactag	aacaagaaag	gaaggtatac
55081	attatttttaa	ataacaaaag	ttaacagagg	cactaataat	aatgacataa	ccacactgga
55141	ggtggagagc	aStgtagata	tcctcattgt	cacagaagtc	agtcaataga	ccgtgtctga
55201	aaactaggaa	acagaaaaaa	acaagacagt	tccttcagg	gaactagccc	caaggtgagg
55261	caggaaactg	atgattttca	ttatagggtta	cccttcata	ctgccatgtt	gacccatgtg
55321	cacaaattac	cttgggtgaag	tttttaattg	ttaaaaacaa	tcattggtgat	tacacactaa
55381	atggctctta	tttaagggtca	tacctggaat	tcctaatttc	tccttgccacc	acaggggcaa
55441	tcctggaatat	ccttttcttg	aggaatattt	tcaccagaaa	tcagatggg	ggcaataacct
55501	ctgccatatac	taagaatcta	aatcaatga	agatcatgtt	caaataatca	ataccttacc
55561	tataagttgc	caatggtaac	atgctatcta	ctccatgaat	gttcctactc	ttgatgtagc
55621	actgacccaa	aaggcatgtc	acagttcccc	catcagacct	ggctgtRcca	gtgtgccact
55681	aatgccttct	caatcacctc	aaagtgtatta	tttcagttta	tctgactcag	agggcatcaa
55741	aatatatctc	ccagatgatg	cttttactac	ctaattgttg	caacttaatc	ctatgaatat
55801	attgtgaagg	gactaagaat	gagcctctgc	tctaattgca	gaattctgcc	cagagtctgt
55861	gcctaccttc	atagttaaaa	aatttttagga	gggacaaata	ccaagtga	catagtgttt
55921	tgaaaactac	tacaaacata	agtaaatttc	actgtaataa	gcttcctaca	gcaactgagt
55981	ggttttctgt	atthttgtcta	aaagcatatg	cattgtctaaa	aactgcctta	gtgttttaaga
56041	cctagatcta	ttcttctctgt	gtattttattt	gaaccagtga	ctggtttatg	ggagttagt
56101	tttctttctg	gattttacgtt	tatggttaggg	gaggttaagg	agaaaaatgt	taacatgtca
56161	catttttacia	gccaaagtta	cctgttggaa	atggggcaaaa	ataacctttt	ttctttcttg
56221	cgggggggcc	aatggtgcct	aaacctcatg	taccttaggc	aacatctcat	tcattctcca
56281	tcctgatgc	ttgctttaga	aaatgaaccc	tgtatgataa	acagtataac	ctttagtctt
56341	ttagtaacta	ttaaatggat	cagcactgca	aaacaccttt	ctacatggcc	catctgtgtg
56401	aggaactcct	ctaacaagat	aacaaaagcc	tgcttttata	ggctcctaag	gaacagacta
56461	atgttactat	gaagttattt	cttacagatt	atactcataa	aacatggcct	gaagagaaca
56521	cgatgaggag	catgagctc	cattttacct	gttctggttc	aagggctatc	tgagttttaa
56581	acttctgaaa	aatttttatct	tccttggtat	catgttttgc	catggaatcc	agttcttctt
56641	caagtgtctc	acctgaaaaa	tcaacgtaac	tattatgaaa	aacaggagta	atccccacaa
56701	cttgacaatt	cacacatgga	gaggggaccc	acttttaact	agatagcttt	ccctatttat
56761	tcactcattc	aagttggacc	atctgaattt	ccaggtactc	catccaactc	tattatatgg

56821	acttccattt	agtgcattct	cttaaaagctt	caaaaataaca	gaatgggtcaa	gggcttagga
56881	ctgcccagca	catcacagga	cacccaacaa	atgtgagccc	ttatcattag	tatcctcagc
56941	tggtaggctc	actcactcag	tcatacaagt	ttcattttctg	gcctggagca	gtgggtcacg
57001	cctgtaatcc	cagtactttg	ggaggccgag	gcgggcagat	cacctgaggt	caggagttca
57061	agaccagcct	ggccaacatg	gtgaaatccc	gtatctacta	aaaatataaa	aattagccag
57121	acgtgggtgt	aggtgcctgt	aatcccagct	actcgggagg	ctgaggcagg	agagtcactt
57181	gaacctggga	ggcagaggct	gcagtgaacc	gagatggtga	cattgcactc	cagcctgggc
57241	gacagagtga	gactccgtct	caaaaaaaaa	aatgttcatt	tccttctcca	cattccttcc
57301	tgggattaca	gccaccctaa	gccactgctg	tccccaacag	acccgtgtct	ctaagtataa
57361	ccattagtct	ttgtaatgta	cgttaaaata	gaactgatat	accttgggtc	agagaagcta
57421	aaataactgc	tttgatgaaa	ctggaaaggc	actgatggtg	ttcacttgca	ccatcaggtc
57481	tgatggagga	agtgtaggat	gccttcagat	tgatgttcca	tcaagtatac	gtggaaagtt
57541	tcagtataac	cgtaggagca	ctgtaaatgc	tggtccctca	ggccctactg	cctcctgcca
57601	agtctcaggt	aagacacagc	tacctccagg	aagcattttt	ctattcaatt	ctccttttat
57661	tttagaaaa	tttggacata	cagaaaagtg	gaaatactat	aatgaaccgc	cacatcatcat
57721	taatcagttt	caacactatc	aagtcagggt	tttcctttct	ctgccacttc	cacctccatt
57781	actctgaagt	aaattccaca	catatcactt	cattcataat	taagttatgt	acccctcgaa
57841	ggcaaactct	ttccttttat	ttaccaat	gaaaaatgagt	ctgttcccaa	gtatcctata
57901	aagatgatta	ctgagttttt	ttaaagtatc	attttgaacc	cattaaacat	atctgatgca
57961	tgtagcgtcc	ttgtggatgt	tcaaactgtc	catcttttgc	aagcaggagc	cttttcatgt
58021	gtcttgagtt	ctgacatggc	cctagtaatc	cttatcctta	atcttttgata	tgacctagtt
58081	cccacattat	atgaacgttt	cctgacctag	ttctggaatc	aaccacctct	ccaaagagcc
58141	tggagtccct	tttagagaga	aatggtacgt	agacacaatc	aacattatct	tcctcctacg
58201	cccaacatct	cagttctcag	taacaccaac	ataattactc	gtttcccttc	cccccaatac
58261	acacacaacc	atctcaaaat	aaaagcaaca	gtctagtaat	aacatgttta	ttcaaaatac
58321	taacactgtt	acattctttt	catctccagg	gcatattcta	ctagagatgt	actgtcctat
58381	gttttgaagt	cacctggaag	agttcttagt	gtggttatat	gactacatca	agagtttttt
58441	acttttgatg	attagggact	gcttttttaa	acttacttta	ctccataatc	ttaaaatact
58501	catacagttc	cacagtcaca	tttacctata	caaggcatat	ttgaagtcca	gctttcatcc
58561	ttgatcctgc	tactctaggc	tctcctttct	cctaaagata	agcattttca	ttatgtatca
58621	tgtttatcgt	ataggcatga	acacagccgc	ggccccctcc	aggcagtcct	cagtgtatgc
58681	acgtgttccc	atggcacctg	tattgtactc	ttatcagtca	ttatatggac	tttaacttcc
58741	ccagatatta	tttgggctcc	tcataagac	tgtgagcatc	tgaccactgg	agtgttgctt
58801	cccattatat	ccctgttatc	aagcacaagg	tcaggcacag	agtaagactc	aaaacatgtt
58861	ttggaagtga	tgactgggtat	gaactacaaa	ccagtaagct	gatgttttca	ttttgagtc
58921	ataaatctaa	ttttgtgggt	gttttgtgta	tKgctcaagg	ctcaaattgt	aaaatttaat
58981	attatgtgac	caaagaaagt	tatacccaga	acctcaattt	cctcaccttc	aaaatggggc
59041	agttttctc	tcattgggtc	gctgtcaoga	ttttaatgag	ctcatgcaca	aacagccctt
59101	tatataaggt	aagtgtctga	taaatgttgg	ctactataat	aaaataagcc	tctaagatac
59161	ttggtcagca	caaagtactac	ccaagagtat	gcaactgaag	taaactgaca	aaattgtgta
59221	tctaaaactg	gcagatgaa	agagaaaact	tttaaggggc	cttctgctgt	cccgacactg
59281	tgctaggcac	tcacactatc	ccgacccgag	aaacMgatct	gcgacccaga	ggaaacttacc
59341	aagcctccag	catcttgtgc	agccctactc	atgggacat	ctggataccc	acccttgtct
59401	ttacagggag	cagaacacac	ctcttatgtg	tcagaaaaca	aagtccagga	agtatatatt
59461	tacctgaggc	aatatctgaa	aattgtatgc	tcagccctcc	aaagtgagtc	ttcctctcag
59521	tacctctctt	ctaggcacat	ggagcccttt	cttccaagta	ttatgtttaa	ccacttcaatg
59581	aatgaagtcc	tgaaactgct	tacccatgct	ccctataatc	tctgagtaat	cttctctttc
59641	cacaacctca	ggcataatct	catcttctgt	ttctattaca	atttcaaatt	ctggaaaaag
59701	gaagtgtgtg	tctggaatta	tatggtccag	atgatctgaa	acaaaaagga	cagcactatt
59761	agtaatcatt	tagttttgaa	gacagtctaa	taatttgctg	tctctaaagt	actatattcc
59821	ctatagttct	ggcatttttag	ataaaagggtc	ataaattaaa	tgcctatatg	gtgacattat
59881	tcagtgatcc	agacttcaca	gccttttttt	ttttttttaca	aaggtgttcc	aggcatgaaa
59941	aatttttaag	tactatacct	ttcctaattt	tacctttaaa	gttgtcctgg	aaatatctgg
60001	gttgacaaag	gcgatgaaac	tgaactgaRa	cttaaaaaaa	agattaccca	cctgggtgtg
60061	cacaagcctg	cttatgtccc	aatctccagt	ctagggtctg	atgctccttg	ctgcagtaat
60121	atgcttttgt	gcatctggag	cacgttttgg	ggcctaaaca	gccacaaaacc	ctgcagagat
60181	gagcaccaga	cttaagctgg	agacacactg	attctcctgt	ttctggggga	ggattctcag
60241	aaggtggctc	atatgagtaa	aaatcgtttt	tcctgggtag	ttgattccta	aaaactaaaa
60301	agaatacag	agaaaagt	tatcttcaaa	caaaaacagca	attcacatat	tttatcctct
60361	gcacgtataa	ctgaaaataa	caacaacaaa	aaagaaatga	aagtttttgc	tttcaggaat
60421	agctttttaa	aatccagaaa	ctagattttcg	tccggtacac	gcaactgagt	tgccctctag
60481	aggtgggttg	agttaatcaa	attaataaga	ctgatcggtta	agaacgactg	ccaaaaatac
60541	gaaaaagcta	ctgggatcca	tctttccaag	acaatttcta	ttatctgaat	taacaccata
60601	cctggtaccc	actgattaaa	agctgggggt	taccaatgcy	cgtgggcaca	gttagaagct

60661	tatgtagcaa	aaatgagcac	atcctggaag	ggccccggag	aaggtgctcc	tggggcagcg
60721	cggagagggg	gctctgaggg	tggggcgcca	gcggtgcttg	cgcgcgtccc	cctgggtcgct
60781	cccggaaatta	acgcgcgcga	cgcgtcggag	gcatggcccc	gtcccgaacc	cgtttgccgg
60841	ctcacctcgc	aggccggcac	agcacggctg	ctcgcggcag	cagaagagga	agatgcagcg
60901	gtggaaggcg	tccggggcgg	caggcagcgg	cgcatacacc	tgcagcagga	aggagagcgg
60961	gcggccgcac	agctcgcagg	ccagggcctg	gggccccggc	agccccggcg	cgcccagcca
61021	tgcgggccgc	ccgcccacct	tgctggggaa	ctgctcgtcg	cgcagtcgcc	acgcccggcg
61081	cgactcggcg	aagcccagct	ccacaggcct	ggccccggcg	gcagccatgc	ggggcgccgg
61141	ctggcggtgg	gcgcagccca	cagctgggtc	ggaaggcgga	aatcgggcgc	cgggccggaa
61201	ggcaagaggg	gggcaccttt	ccggaggaca	ggaggcgga	acgcgtctga	cgggagcggg
61261	tgcaggacca	atgcgagggg	acggggcaga	ggaaacctct	cggcatcagc	ccgcccctg
61321	gcgcctctgc	ctccgagccg	ctttcctggg	gcctccgggt	gctctgggat	ggttctggtc
61381	tttgggagag	tggcagctgg	tgacggcgct	ccgctcacct	ctgcacatgt	cttgctgtgg
61441	gcctgcgggt	ggccgcccagg	gaggcagagc	cctcccKcaa	accttcccctg	ctgggtgtcca
61501	cctcaggggtg	tgggaaacct	gtgcgctggc	cgagtgttaa	ccaagagtga	gcagtgaag
61561	acaaatgaag	gttgaacagg	taaagtggag	accctacagc	ggaaaccaag	aatcctgtgt
61621	gcctgagagt	aatgaagaag	cctctgcaga	agagtctttt	ctgtcagctc	taaggtctct
61681	gttttaatat	tagtgctggc	ttgctgtacc	tgaattccaa	gggaggagtg	tataatgagg
61741	catggccaac	ccccacttcc	catcattgcc	tgaactagtt	tttcagggtta	acttcagaat
61801	gcctctgggc	aagcagaggg	tccatcagtc	ggttggaggg	tttagaattt	tactgttggg
61861	ttgcaaagggt	ctgaaagaaa	catgtaccac	ctgttctcct	taaggagttc	tacttaggag
61921	gtttcatttta	cataacaaga	ccgtggttgt	cagccaggtc	tccaccgcga	taacctgtta
61981	tgccacaatc	caaaccacca	ttctgttaac	tcaagatggg	atataagttt	ctgaaccca
62041	tttggggcct	cagcaaaatc	actctgggtc	tccccatgt	gcagtgttaat	aaatttgtat
62101	gccccttctc	caattaatgt	gccttttgct	agttgacttt	tcagtgaacc	ttcagaggac
62161	aaaaaggaag	ctttcccttg	gclactacag	tggcittaat	ggaagtaaa	tcacaaaca
62221	cattttatttt	tgacaaaatc	acagttagt	tggcagtata	tttgtttgtt	tttgttttat
62281	ttgagatgga	gtctcgtctc	gtcgcgcagg	ctggagagca	atgggtgcgat	ctcggtcac
62341	tgaacctcc	gcctcccag	ttcaagcatt	ctcctgcctc	agcctcccga	gcagctgggt
62401	agcagtatat	ttgtaattgt	acataaaata	atgtctgttt	ttaaataaac	atttacattg
62461	taaaccaaaa	gttaactctc	catcaatttt	ttttttctct	ttttgagatg	gcacatctgt
62521	ctgtcaccca	ggctggagga	gtgcagtggt	atgatctcgg	ctcactgcaa	ccttcgcctc
62581	ccaggcaagc	gatcctcttg	cctcagcctc	ctgagcagct	ggaattatag	gtgtgtgcca
62641	atacaaccag	ctaattttcg	tatctttggt	agagacagag	tttcaccatg	ttggccaggc
62701	ttgtctcgaa	ttctgacct	cagttgatct	gtccgtctca	gcttcccaa	gtgtgggat
62761	tacaggcatg	agccaccaca	tccggcatcc	atcaatttag	aaagtattat	tcaccaagat
62821	taaggttgca	cccgtgacac	agcctcagaa	ggcctgatg	accatgtgcc	cttggtggtc
62881	agggtacagc	ttgcttttat	acatgttagg	gagacatgag	acatcaatca	gtatgtgtaa
62941	gatgtacttt	agtcaggtaa	agcgggactt	gaggtgaggg	cttcagRtc	atgagttagt
63001	aagaggcaaa	agatgcattt	tttttgagtc	cttgatcagc	cttccactga	atacacaatt
63061	tagtctggct	cagtgaatta	tcatttttSa	gtaaacata	ggggaggggg	agcaattaga
63121	tatgcatttg	tctcagggcg	accttaaagg	gataactttg	agttctgtct	gtcctttatc
63181	cacaaggaat	ttccttggtg	gcaaatttta	agggaggtac	gtagcctctt	atcctggcag
63241	ctatcttatt	taggaataga	atgggaggga	ggtttgcctg	acatagtttc	cagcttgact
63301	ttaccctttg	gtttagtgat	ttgtggtcc	tgagttttat	tttcctttca	cagaaattat
63361	accgtaaaag	taattgaaga	aaatcacctc	tttccccttc	cctcaactag	gccttgacca
63421	ttttaaataa	aatcaggatt	tgctgaagg	caacaaattt	aaccaagttc	agttaaaact
63481	taactctgaa	tctgtatgtc	cctggggtct	tttcagtgga	gagatgtcta	agcatcattc
63541	caagcttttc	tatactaact	ggcctatttg	tatgtttctt	atttttaggat	tccttttgtt
63601	catgtgtatt	tttattaggg	aatcacccat	ttcctgtagg	tttcagggtt	aatatatttt
63661	tcttatttga	atttttaatt	atcctctcta	gttttctaca	tatttttttt	ttctttgaga
63721	cgagggtctca	tatgttgccc	cggttggcct	caaactcctg	ggcttaagca	atcctccac
63781	cttggcctcc	caaagtgtcg	ggattatagg	tgtgaattac	tgtaccagc	ctagaatcct
63841	tagtccctata	tacttttgct	gtttttttca	ttgccaactt	gaaataataa	aaagggctag
63901	aatccctataa	acaaaaata	aaataaggcc	ccccaacctt	ctgaatggac	ttcctcctc
63961	gacacagatc	ttttacaatt	taacctgtat	gaaccccaaa	aattggagac	aggtctcagt
64021	Kaatttagaa	aatttatctt	gcgaaggcca	aggacacacg	attatgacag	cctcaggagg
64081	tcctgacgac	atgtacctaa	gatagtcaga	gcacagggtg	gttttatata	ttttaggagg
64141	ataaagaca	tcaatcaaca	tatgtaaaa	gaatattggt	caggaaagg	gggacaactc
64201	aaagaaaaaa	tgggacaact	cgaagtgggg	aggggcttcc	aggtcacagg	taggtgagag
64261	acaaatggtt	gcactctttt	gagtttctga	ttcaccttcc	taaaagaggc	agtcagacat
64321	gcattttatct	cagtgagcag	agggatgact	gaatggaatg	ggaagcaggt	ttgccttaag
64381	cagttcccag	cttgactttt	cccatttagct	tagtgattat	gggattccaa	ggtaatttcc
64441	tttcacattt	ccccctttt	cttttttaaa	atattttgga	gaaagcattt	ttgaagaaaa

64501	taagtttctg	ttcccaggtt	ttatctgctc	tctcatggct	aggatgggtt	tttcctagaa
64561	gggtaggtcc	tgagttatta	ggaaagctca	tttttagaag	gttgtgaagt	ctaataatcct
64621	atcaagagaa	aatttgggga	ggaaaggaga	acaataagaa	caatcttgga	aaattgatct
64681	aggccacatt	actctgaagt	ccatacatca	gtaagcaggt	atgaaagtgg	cttatgtatg
64741	taaatagggt	cccattttt	tcttctgaag	tttaagttgt	ctacttcagt	tcacagggtc
64801	tcacgaaagt	tagttttaag	tgacttagtt	tagtgacagc	ttagttttta	gtgactccaa
64861	attaggaaaa	atggggaaaa	aaagaaggaa	aaaaattgaa	aacattat	tgaagacttg
64921	tagcccacaa	aaattagaat	ttggtccaaa	ctgtagaaaa	tgataaaaa	tgaaaaacat
64981	taggcaagac	tagaatctaa	caactgggtg	actatagttt	tccagtcctc	agtttcccat
65041	ttatactaaa	gacaaatcat	gataggtttg	ccttattata	tttggccgaa	ttatttgtat
65101	acagtgcagc	aagaataatt	attttttaac	attggctttt	aaattggctt	tgatggaact
65161	ttgttccata	gagggtat	cagataagac	tttttaaaag	ctgagcccag	ccatggattt
65221	gtgccatcaa	atacctgtga	gtttggtgat	cctctcctct	tgaggttcca	agataaactt
65281	gaggctcctg	ggcctgtcag	aaagtgcacat	tctttactta	ccacaggcca	ggaaccctgt
65341	acaggagctg	tgtagacaaa	gttatgagga	cggtttttcc	aaggggggtt	aattggctcc
65401	gtaagtcaag	taaaagcatt	ggttaacaac	cagtttcccc	aattgtgtcc	tgttacaaat
65461	gaaaacagat	tattagtgc	cttatgcaaa	taactgtatt	gtcataagtt	aagaataactt
65521	acagtttcca	aattctggag	aaatcggtta	gagagaacca	aatgtgctcc	aaattttgtt
65581	cataggagta	tgtgttactc	aattgttaaa	agctgcagat	agcctgacca	acatggtgaa
65641	accctgtctc	tactaaaaat	acaaaaaata	gccaggcatg	gtggYgcga	cctgtactcc
65701	gcgctacttg	ggaggctgag	gcaggagaa	cacttgaact	tgaggagtg	aggttgacgt
65761	gagccgagat	cgacccgctg	cactccagcc	tgggcaacag	agcaagactc	catctccaaa
65821	aaaaattatc	ttgtttttat	caatctttct	taaatgtata	gctcacattt	atttcaatgt
65881	ttaaaaatgag	aaatat	ggtctttatt	tagacgtttg	ttgatgtttt	tgtgaccaga
65941	aatatgtttac	agaactta	ctcttgttta	tatcaatcat	cccatggcaa	attgggttta
66001	ttaccagtag	tggtgcttaa	agtcacagtt	tccaagaatc	tatccatgat	aaatgaggac
66061	ttactggact	taatagtact	gaattgtaca	cataaaaaatg	gttaagatga	taaattttat
66121	gttatgtgta	ttttaccaca	atttttaaaa	atgaaaaaaa	gaaatcctgg	ggaatgat
66181	gttcaaagaa	ggctttgaaa	agctctgata	tattccagaa	atctagacag	ctacacaggg
66241	gcagggcttt	gtgctgggcc	atgaaagacc	tgaggcagct	gtaatgtttt	acatttgact
66301	ctggggcttct	gtgccatcag	aaagtgaagg	ctaaagcaga	cttggaactgc	cagagtgttg
66361	aagatgtgtc	ccagcacata	caaagccact	ctacaaaggc	tgaggagacat	agggttcaagg
66421	catttaagga	aatctctatc	taatcattag	ctgaccactc	agcacactaa	gcaaagactg
66481	aaagacacat	gacaaagaat	acagacttca	tagaattagt	cagggaaagt	cactaaaaaa
66541	ttcaacagca	acaacaacaa	aacctttgag	agggaagggt	ctggcttaca	gttaccacat
66601	cctagtcttt	aaaatattca	gttttgagca	aaaatatgag	atacacaag	aaacaagc
66661	ccatgcagag	gaaattctga	gaccagctcg	atcggggaga	ccctaacc	gtggcactag
66721	gggatattaa	agacacacac	acacaaatat	agaggtgtga	agtgggaaat	caggggtctc
66781	acagccttca	gtgctgagag	ctccaaacag	agatttacc	acatat	taacagcaag
66841	ccagtcattg	gtattgtttc	tatagatatt	aagttaacta	aaagtatccc	ttatgggaaa
66901	cgaagggatg	tgccgaatta	aaggaatagg	ttgggctagt	taactgcaac	aggagactgt
66961	ccttaaggca	cagatcgctc	atgctattgt	ttgtggctta	agaatgcctt	taagcagttt
67021	tccgccctgg	gcggaccagg	tgttccttgc	cctcattccg	gtaaaccag	agccttccag
67081	cctgggtgtt	atggccatca	ggaacatgtc	acagtgtgtc	agagattttg	tttatggcca
67141	gttttggggc	cagttttatg	ccagattttg	gggggcttgt	tccaacagg	aaaataaaga
67201	agtcaatgga	aattctccat	gaggaggccc	agatgttata	cttcctaggc	aaagatttta
67261	aattagctat	tataaatatg	ttcagagaa	taaaggaaac	catttctaaa	gactaaagga
67321	atgtataaga	atatctcacc	aaatagagaa	tactaatgaa	gagataagaa	ttagaagaac
67381	caagtaagaa	atgctagagt	tgaaaattac	aaaactgaaa	tacaaaattc	gctagaggca
67441	ctcaacagca	gattagagat	tgcaaaagaa	agaaaacagca	aacttgatga	taggtacatt
67501	gagattaccc	aatcagagga	gcagaaagaa	aaaagaataa	agagaaatga	acagagcctt
67561	agaaatctgt	gggataccat	caaacaatc	tggaagtccc	agaggagaga	acaaggggga
67621	aaaaaggcag	aaaggctatt	tggaagaaata	atagccaaaa	ttttccatta	agcttttcat
67681	ttgattaaaa	tttatatat	ttaaaaagtt	taagtgtatc	ctactcatat	tttttaagaa
67741	ggtccatacc	tagacacatc	atagtcacaa	tgccaaaaaa	gagagagaga	atcctgaaag
67801	caccaagaga	aaaacaattc	atcttgtata	agagatcatc	agtaagatta	acagctggct
67861	tcttgtaaaa	aactcacaga	gaacaaaatg	ctgcagaatg	acatacttgg	gaaggcagtg
67921	gggtggagaa	gcatgccaac	taagaattgt	atatccagct	aaactagcct	tcaaaaacca
67981	aagaagaaaa	ataaaaaaaa	aatgagatat	tcccaaat	ataataaaca	aataacactg
68041	agataatctg	tcattcatag	atctgcctta	tgagaataac	taaagggcct	cctctttcag
68101	gctgaaataa	agaacatta	cacataactt	gaatctacag	aagaaaaaag	agcacggtta
68161	aaagtaacta	cataggtata	taaaaatgat	agtataaatg	tatttttgtt	tgcttgtaac
68221	tcattctctc	tgattttaa	gcaactgca	ggaacaata	ccaataaaaa	tgttttgaca
68281	atgttatatg	aaaattta	ttgcatgatg	gtaataccac	aaataagagg	gatgggaata

68341	gagctatatt	ggagtaaagt	ttttgtgtat	gatttgaatt	aagttggtat	taatctgaac
68401	tagattgttt	taaattaaga	tgataactga	aattaaatgt	taattgaaag	gcagtcacta
68461	agaaaataga	gtaaaagaaa	caaataaata	aatggcact	taaaaatatc	taacagtaaa
68521	gaaagcaata	atggagagca	aatgacttg	aaacatacag	aaaataaaga	gcaaaatgac
68581	aggcacaat	cctaccatat	cagtaattaa	atgtaaatga	attaaacatt	ctaataaaaa
68641	aggcagagat	tggcaaaatg	gattaaaaca	caaaacacag	acagaaaacag	caaccattct
68701	ccaattatat	gaggagacct	gctttacatt	caggaaaaca	aataggttga	aagaaagggg
68761	atggaaataa	atatactctg	caaacagtac	ccaaagacag	ctgaagtggc	tatataaaca
68821	ccagacaaaa	tagactttta	tgacgaatt	attagagaca	aagacatttt	ataatcaaag
68881	ggtcaatgta	tcaggaagac	ataacaatta	taaacatata	tacctctgac	agtagagcac
68941	caaagtaaat	gaaacaaaaa	ctaacttaaa	aagagaaata	cacaggtcaa	caataaatgt
69001	tggagacttt	aatatccac	tttcaataat	gaacagatct	aggcagaaga	tcaacaaaaa
69061	atagaagact	taacactatg	aaccaaatag	acctagcaga	catctgtata	atactgcac
69121	caacaactgc	acaatatata	ttcttcaaat	acacatgaaa	gattctctag	gatagaccat
69181	gtgttaggcc	ctaaaaacaag	tctcaataag	tttaaaagga	ctgaagtcat	acaaagtatg
69241	tcttccaacc	acaataaaaat	gaaattagaa	atccataaaa	gaagaaaaatg	tgagaaattc
69301	acaaataagt	agtaatgaaa	caaaataacc	tatgggataa	aaaggaaatc	acaggggaaa
69361	ttagaaaata	tgtagagatt	aatgaaaata	caccatgtca	aaatttatgg	gacacaggac
69421	agtgttgaaa	aggaaattta	cagctgtaaa	cacctatatt	taaaaagaaa	gatttttagtt
69481	caataatcta	aacttctacc	attagaaatg	aaaaaggaaa	tgcaacaat	ctaaatcaag
69541	cagaaagtag	aaaatagcag	agattaggg	ggaataaat	ggaagacca	aaaaacagag
69601	aaagtaataa	aatcaaaagt	tggttattgc	cgaaccagc	ttggctgggg	agaccctaac
69661	ccagtggcac	tagaggaatt	aaagacactc	acacagaatt	atagaggtgt	ggagtggaaa
69721	atcaggggtc	tcacagcctt	cagagctgaa	agcctcaaac	agatttacc	acatatttat
69781	tgacagcaag	ccagtgataa	gcagtgtgtc	tatagattat	agattaacta	aaagtatttc
69841	ttatgggaaa	caaagggatg	gcccgaataa	aagggtggg	tctggctagt	tatctgcac
69901	aggagcatgt	ccttaaggga	cagatcgctc	atgctatttt	ttgtggttta	agaacacctt
69961	taagtggttt	tccgccttg	gtgttacagg	tgctccttgc	cctcattccg	ataaaccac
70021	aatcttccag	catgggtgtc	atgaccatca	caaacatgtc	acagtgtgtc	agagattttg
70081	tttttgcca	gttttgcggc	caatttatgg	ccagattttg	ggggcctagt	cccaacgtgt
70141	cccccttctt	tgatttgcaa	agtataaaaa	gcaaaaggcag	ttttgtcacg	gtgagctact
70201	tcttgagga	gtcaggatcY	gcactgtcag	actatacaaa	aacaacatag	attaaaagca
70261	caatcatcat	cgaaatcaca	gagcttccaa	gtgttttcat	ccattttaat	gggttgctag
70321	ctKcttatct	gtctgcagct	cctttaagca	ctccttttcc	tggcattaag	gtcaggtgtg
70381	cctaggatgc	tttattttgt	cttttaattt	tgcaatatcc	aaaaacaagc	ttgtagagtg
70441	tccttctaga	tgctttttta	atctcttccc	aaattttgat	ctaatttaaga	gctatttaata
70501	atttccacaa	atccttattt	aagctcctag	agtgggccat	atcatttgag	gttgaggtgc
70561	cactataccg	ccatggttcc	agatgatagg	aactcttgcc	atgtcctatc	atttctacca
70621	tctgaccatt	ttgttcagac	cagctgaaca	tagtgtggct	gtggcatgca	gactgagagg
70681	tgcaattcaa	gccaaacatc	cccttagggg	accaatcaat	aatgattcca	taggaagcat
70741	tgtgcagcac	ctctgcctgt	tctgcaatgc	agtcttccca	aacaagtaaca	ttcatttttt
70801	ctaactgggt	ccaatcctgt	ttacaaaatg	gtttttgagg	gcagtatgcc	ttattatag
70861	gagcagattt	attatggtaa	atactgagat	cagaaagcat	gtgtaagtgt	gtcatagagt
70921	gattacatcc	aggcattatt	gccagccaag	attgataaMt	atgcccaata	agtacagttg
70981	ttctctgtgt	cagcacttgt	tgaaggaata	ctcatggcaa	tggtgatcac	cgctgtcata
71041	gctaccatta	aattactcac	tgtgactgg	tgtcctgctt	tcctcagggt	ttcttctgcc
71101	aactgtgaca	gcttcttgat	ctgtcccca	gtgggtggct	gtgtttgatg	gggtgtgtct
71161	gtgacagttg	gggtcctcct	cagcatcagt	cttgagatgg	ctgcaaccag	ggggtcctca
71221	ggatcctcct	ggaatctctt	cctcagcatc	tggctcatga	taaggtttca	gggtatcttga
71281	tggtatccaa	atcagctctt	gattttggcc	tgagaaaaca	caagcataac	ctctacccca
71341	agttattatt	ttacctattt	cccaaccttt	tgttatgcaa	agtgagaaaa	gggtgtgcaca
71401	catacatgta	acgtgtgaca	tccatttgcc	aaagataatt	agggtccaat	cctcgaggat
71461	taactcctcc	tgcaaaagat	gaggaatgca	ccattttggca	agttgggcat	cactggatca
71521	taacttttagc	ttctttccag	gtaatgctgt	atctgtgttt	cagaccagag	gcattaacat
71581	gggttaaatt	tgaaaaatgt	ctagcagtag	atattgcttt	agcaactagg	tgatcagcca
71641	tttgattccc	ttcagtcaaa	ggttcggaa	gaggtgtatg	agccctaagt	tgagttagtg
71701	aaaaaggggtg	cattctactc	ctaactgctg	tttgcaactg	ggtaaataaa	gtcatttagtg
71761	tcatctgttt	gaaattgtaa	ctgagcattt	tcaactaact	gtgtggaatg	aaccacgtat
71821	gaagaatcag	aatcatattt	aatagacata	ttaaaagcag	tcaatacctc	aattacagct
71881	acaagtctctg	ctgttagagc	tgaagtatag	ggtgtctgga	aaactttaac	ttttgagcca
71941	gaataagaag	ctttaccatt	actagaccca	tctgtgaaac	aatgaaaaatg	cttagcaggc
72001	tgcaagttgt	ttactgcagg	aattataaat	gtaaaccatt	cacagtcttg	cttagctaaa
72061	gtgatagtaa	agaaacagtc	ctttaaatct	gtgacaatta	aaggccaatt	ttttggaatt
72121	atagcaggag	aaggcaatcc	aggaggtgt	aatgtcccca	taggttgtat	aactgaattg

72181	atagctatta	agtcagttaa	cattctccat	ttacccgatt	ttttcttaat	tatgaaaact
72241	ggagaattcc	aaggggaaaa	tgctggcgct	atgtgcccat	tttctaattg	ttcagtaact
72301	aattcctcta	aagcctccag	tttctcttta	cttagcggcc	attgttctat	ccaaattggc
72361	ttatctgtta	agcattttta	agggtgaggt	tctggaggct	taacagtggc	cgccatcaaa
72421	aatgatatcc	tacaccttgg	cgggaacttt	gtctttacgc	ttgaagcggg	tccttcaaac
72481	cttgcaattt	ttttcttagt	cccataccag	ggacatgcc	catttcacgc	atcatagttt
72541	gactttgagg	gctatataat	tgttctggaa	ttagaacttg	tgctcccat	tggttgtaata
72601	aatctcttcc	ccataaattt	glaggtacag	aagttataat	tggttgtaata	gtcccagggt
72661	gtccatcggg	ccttcacaat	gcaaaatata	actactttga	tatacttcag	gggctttacc
72721	aactccaact	atgttaaatt	gagcgggtgg	aattggccac	gcaaacagcc	agtgtgttag
72781	agaaatgatt	gaaatgtctg	ctcctgtatc	taccaaacct	ttaaatttct	ttccctgaat
72841	agttattttca	caggtaagat	gtttatcagt	aatttgattt	acccaataag	cttctttgcc
72901	ttgtttattt	gtgctttcaa	atcctcctgt	tcgtttaatt	taacttttcc	gtattttctac
72961	atacggcaca	atcaggagct	gtgctataag	ctctcctggc	tctgctttcc	agggaaacaga
73021	agtggatata	acaatttgaa	tttccacatt	gtaatctgaa	tcagtgactc	ctgtttgtac
73081	ttgtatccct	tttaaattta	aactagacct	tactagaagt	aatcctatca	tcctctgtgg
73141	caatgggtcca	cagactcctg	ttgggaactt	ttgtgggggt	tccccaggca	gaaggctcac
73201	agcattttgtg	cagcataaat	ctactgcagc	agtaccagct	gtggcagggg	acaaacattg
73261	taaattcccat	gaagtatatg	gggtattagt	gcattcattt	agagctgagg	gctcttaggt
73321	ggaagatggt	aataaaactgt	ccagggtccca	tggttcagca	gagcattttc	cattgtctggg
73381	gcccttaggg	atctccagga	ttatctccag	agtgccata	atggacagta	caaaaStggg
73441	aagagtttcc	aagaaagcaa	atccccatc	atccaccaca	aaacacactc	tcagcttatt
73501	catcaatgtg	gcaactgggt	ccttaatgtg	tggcaggcag	aggtaatgac	agtgagcaaa
73561	tagcccaggg	attacttgag	ctaaaagcaa	caaaggaaacc	aaacgcttcg	actttctgac
73621	caaactcagt	tcccaactga	agctgggtctc	atgcaagtgt	gctggcctga	gtgatctcac
73681	atcctgtgaa	ggctctgcca	aggagggatg	gtggctaggg	tcaggcctcc	ttcaaggggg
73741	tcccgatcc	taccagaggc	taaggtggga	ggatcgcttg	agcccagaag	tttcagagca
73801	gcctggacaa	catggagaga	ccctgtctct	gcagaaaaaa	aattagcaga	gcatgggtggc
73861	aggagtgttt	aatgccagct	actcaagaag	cagaagtggg	aggattgcct	gagcacagaa
73921	gtaccaacct	gcagtgaact	atgattgtac	cctgacactc	cagcctgggc	aagagagtga
73981	gaacctgtct	caaataaaga	aaggaagaaa	ggaagaaagg	aaggaagaaa	ggaagaaagg
74041	aaagaaggaa	ggaaggaaaa	ggaaaaggaa	agggaaagaa	aaggaaggga	agcaaaaggga
74101	ggaagggaag	aaggaaggag	aaagaagagc	aaacagagaa	aagaaagaaa	agaaagaggt
74161	gctttggaga	aaccacacaga	agacttttac	actttagtgt	tgttttgaat	acttaaatcc
74221	aacatacaaa	ggaatagaaa	ttatatctct	cacaattatg	atttttcttt	tctttctccc
74281	tttttttttt	tttttgagac	agagtcttgc	ttcttcgccc	aaactggagt	gcatttggtc
74341	gaccttggtc	cactgcaact	tctgcctcct	ggtttcaagt	aattctccta	cctcagcttc
74401	cccagtagct	gggattaaag	gcattgtgcca	ccacaccggg	caaatttttg	tatttttagt
74461	acagacaggg	tttcttcatg	ttggtcaggc	tggtctcgaa	ctcccaacct	cagggtgatcc
74521	gcctgcctcg	gcctcccaaa	gtgctgggat	ttaaaggcat	gagccaccgt	gcccggccag
74581	gtcttagatt	tttttagaaag	tttattttgc	caagttggag	gatgcgtgac	tgtagtgcat
74641	cctcaggagg	tcttgacaac	atgtgcccac	gggtggttggg	gcacagcttg	gttttatata
74701	ttctaggggtg	acatgagaca	tcaatcaata	tgtgaaagat	gtacgttggt	tcagtccaga
74761	aagggtgagac	aacttgaaag	gaaggccaaa	cagggggctt	ccaggtcata	ggtagaaaag
74821	agaccaattg	tttcattctt	ttgagctgct	gattaccctc	tccaaatgag	gcaatcagat
74881	atgcatttat	gagcagacag	gtggctttgg	atagaaatggg	aggcaggttt	gcccctcagca
74941	gttcccagct	tgacttttcc	ctttagctta	gtgattttgg	gtccccaaga	ttgatttttc
75001	ttttgttaagg	tctaacatgg	tttcttatga	gcattaatta	ttcattgtgt	attttattac
75061	aaaaataagg	cacagatttt	ttaaaaaaca	tcaatttcat	gactagtatt	atacacataa
75121	ttacactgaa	gttcaactaa	atttggaaac	attccagagt	ttgggtttct	aataattctt
75181	tggtgattctt	tagaagctaa	aatattttaa	caaagcaaca	tctaaaatca	cctgtagaat
75241	gtcctgccat	ttttgtttct	ctagtttcct	cattttctgc	aaagcctcgc	tgaggaaatt
75301	gactctgaat	atcctttttac	actcttctgt	tttagaaagc	attgtggtga	aacattgaat
75361	cataatggtc	acaagtctctg	ttcacattct	ttctttcttt	gaatattttt	tcccagtggc
75421	caatatttga	ttctgttgta	ttatgggttaa	aaggtaggca	tgagaacaaa	ataaagacaa
75481	gaggtctttt	gaataagtga	tccagtcaca	atgaatcaat	ttgccattgg	aacataattt
75541	tacgtcactc	ttctgaaaaat	atttagccat	gaattgaaag	agagtctgta	agattatttt
75601	tttcctgttc	taagggtgaac	agcatttttag	agaatgaacc	acaaccacag	cacaagaaaa
75661	aaatctgata	aataagttta	cacatatgtg	ttactactgt	aacataaaaac	atgtaaagag
75721	catttggttt	gatttatata	tcagtctgta	ttgtttaatt	ttttgtgtca	taaatgtctc
75781	tattttaaaag	acaggactat	ttaacagtgt	aaattactag	taattcatgg	tataaataat
75841	taaacaagga	agtgttcaaa	aatataacaa	tgtttttaaat	aagcccattt	tgtgcttctg
75901	taacagaata	cctgaggctg	ggtaatttat	aagtaaaaaa	cgttcatttg	gttcacaata
75961	ctgggtggtg	gaatgtctga	gattggggcag	aggcatctgg	tggggcctca	gtctttttca

76021	cctcatggtg	gaaagtggaa	ggggagcaag	ggagtgcacc	agagatcaca	tagcaagagc
76081	aaaagcaaca	gtgaagccaa	ggaagtcaga	ctcttttttaa	ctacctactc	ttgcaggaat
76141	taatccattc	ctgtgagagc	agaactcact	cacctctgtg	gaggacaata	atctattcat
76201	gaaggatcca	tcccatgacc	caaacacctt	ccactaggcc	ccacctcccc	acactgccac
76261	attgggggtc	aaatttcaac	atgagatttt	gcaggggaca	accacatcga	aaccttagta
76321	atttgtagca	tagttaaatt	ctttttcaca	tgatgtattc	tgtgctggga	tactccacat
76381	cctgaatatt	ttaatttaat	ttgaatagag	tttgattttac	ccatttttgc	gtaaaattcc
76441	gtgtgttttg	acaaatgcat	agttgcagg	atccattatt	aaagaatcat	atggaatgct
76501	tcaaattccc	accccatgca	gccaatggca	ttcccatctg	tgcagtttgc	ctttctccaga
76561	atctcattaa	atgagggtcac	actgtgtgtg	ttctcctcag	actgtcttct	tccactcagt
76621	aatgtgcatg	caagattcac	tcatgtcttt	gtgtgagttg	atagcttgtt	ctttcttatg
76681	gctaaatagt	attccattgc	atgaatgtac	cacaatttgg	ttatgcattt	tagggagcaa
76741	aaccttcttc	ttctaacttt	gttccagggt	tggagacctt	caaattaact	gacaatagat
76801	acattagtag	gagaggcaat	acttggcttc	ttattccaca	agtatcattg	tgggacaaaa
76861	ttcatcagat	ggcaggatct	agtttacaaa	gaggtgaaaa	tagcccaaaa	cgagaaaaca
76921	gactagaatc	tgataaccca	caagggctat	agttttcctt	tttaaaaaaa	tttttttttg
76981	agacagagtc	tggtctgtgc	gccaggctg	gagtgcaatg	gtgcaatccc	agctcactgc
77041	aacctctacc	tcttgggtgg	aagcgacct	ccctcctcag	cctcctgatt	agctgggact
77101	acaggcacat	gccatcatgc	ccagctaatt	tttgcagttt	tagtagaggt	ggggttttga
77161	actcctggcc	tccaaaagtg	ctggaattat	aggcttatgc	caccatgcac	ggctgagtta
77221	tacttttcca	ttgaaacata	aaatttctct	ctgtagtaac	catcattttt	gatcatagat
77281	aatcaatgtg	agattattct	tgttttataa	ataagtctag	tttcgttaga	ttttgcctga
77341	gtattttatg	aagtgcagca	agaacaggag	gtgaccacgt	aggtgctttc	aagcttcttc
77401	gctggaagtt	ttcatacaga	atctcagatt	ggacttttaa	aggccttatt	gaggctaaaa
77461	gccaaagcaa	gaacatactt	tcaaatttca	gctgcagtca	ttgtagcttt	atgtgaattc
77521	ctctcttctt	gacgccccca	aatatcccc	aaattcctgg	gcctagcagg	aaattacctt
77581	ctttactaac	ctgtaaggct	gtgaaccgtg	taatgtaggt	accaggctgg	cttttctcaa
77641	agtgtcttgt	aagcattggc	ttcataaaag	tcagccttag	ttccttaaaa	ttgtgtgtca
77701	taactgatct	cagggtatact	attcctaaat	atgatattcc	agtaaaagcc	tgataatata
77761	atcaaagttt	ccaattatgt	cttgcataaa	ggtgaacaga	tgcttatttg	actttctgta
77821	acaactatat	tgtoatgaaa	ataagagtag	tcagtaagat	tttcaaaatt	ctggagaaat
77881	caggcaggga	aaaaaagatt	tttttttttt	ccccaccag	cctcatttct	gtttacaaaa
77941	gtataatcta	ctaaattttt	gtgagttata	gttagcttaa	gagaaagaga	tttcttaaat
78001	ccagaaacta	gaacattaaa	gaaccagcag	tactccaaaa	aagctataaa	attataatca
78061	attttcatca	cttcttcag	tgccatgtaa	tcaattccag	tcttgctgga	tcttgggtta
78121	gcagtgtcac	gaacccatcg	atttctcaac	cagacttctg	gagatcttca	ctgagctcaag
78181	tgtatggtct	taaagttatt	taagcaatat	catcagaagc	ctacaaccag	agtacctgtc
78241	ataggctttt	tcgtgagctc	cagagcgagt	cctgtgttgg	agacaaacat	tctgacctgt
78301	agctgattgc	aggagctttc	aggaaagtat	cagggggaaa	tagtatctaa	atgccaaaga
78361	gtatgaaatg	gctgtgatga	aaggtcagat	gagagttcat	tataaccaca	ctgacaagga
78421	tattcgattt	ttttgggtgc	atacaacggt	taaaataata	attgaaaatta	tgactcaaaa
78481	cagtataccg	gcacatagca	tggaataagg	ggacattgac	aaatttccag	taattttata
78541	caattttctg	aaacataaca	ttttatccat	acaaatataa	cccagggaag	gttaggtatc
78601	tctttttatt	ttatattttg	tatggttttt	cttataaaaa	atacatccta	ctttacctgc
78661	aaaacatgcc	ctgctttttg	catacttttc	atagagtgtc	ttctagtttt	taatacttag
78721	taatctctat	tttccagaga	aactaggaag	aagacaattt	taaaactgtc	tacattagca
78781	ttctatagta	gatttagaaaa	tgtatgagta	taccatctcc	caacatctag	agggatgtgt
78841	ttcctcattg	tacaatttct	cagtgtggta	gacaaaaata	cgtttattaa	cggggcaaaa
78901	tatctttact	ctctctgtaa	aaacaagaag	ccaaaagtat	ataaaactga	attacttatg
78961	ttcagtaatt	aatgttttag	tatcgtatct	tattttaaatt	gatctagata	ttgaatgcaa
79021	atcttttact	tagctttaact	ttaaaggttaa	aaattaccaa	aagtattttg	gaaactatta
79081	ttaggcagat	ttactgtaaa	aaattattat	tgaataaatg	cttttaataa	gaatgacaat
79141	tagaatcaaa	tctataagct	ttaaagattt	aaggatctag	taagtataat	attagcttat
79201	ttgagtagaa	ctcaagcaga	atagaaaatt	gttttatatt	taatagtgat	aactctgaag
79261	acatagtgtg	tttattacac	caaaaatata	aaattactct	tatttaacta	agttttctcg
79321	aaattgtgtt	aacttgaaaa	acatttggat	cagttcctat	atttatttga	gtttggggaa
79381	tattttattta	taaatgcttg	gttttttttc	caagccaagt	tagaatagag	cacttttaga
79441	agattttata	agtgaatttt	gcaatgctct	ctggagtga	gaaaaatcac	atatacataa
79501	catacattaa	tatacataca	aacacaaata	gaggtctcat	agctttcatc	ctgaaatatt
79561	agccatagaa	caggcataaa	tattctgatg	gttaatttta	gacatctgct	tgattgtatta
79621	agagatacta	acatagctgg	aaaagcacaa	tttctgggca	caagtgtgag	gggttttctg
79681	caagacactg	agataaggaa	gatccaccct	gacccaatgt	agataggcac	tgatatgggt
79741	tggtgtgtgc	cccacccaaa	tctcatcttg	aattgtagtt	cctataatcc	ctacatgttg
79801	agggatggac	cctgtggtag	gtgattgaat	catggtggtg	gttactgcc	tgctgttctc

79861	atgatatgtga	gtgagctctg	atgctctgat	ggttctacaa	ggggcttttc	cccttttgct
79921	cagcacttct	ccttacagct	gccatgtgaa	gaaggactct	ttgcttcttc	ttttgccatg
79981	attgtgaggc	cctccagcc	atatggaact	gccagcccat	taaacctctt	tgttcttttt
80041	tttttttttt	tttttttttt	ttttttttga	gatggagtct	ggctctgtct	cccaggctgg
80101	agtgcagtgg	cgtgatcttg	gctcactgca	agctctgcct	cctgggttca	cgccattctc
80161	ctgcctcagc	ctcccgagta	gctgggacta	caggtgcctg	ccaccacccc	tggttaattt
80221	tttgatattt	ttgtagagac	ggggtttcac	tgtgttagcc	aggatggtct	caatctcctg
80281	acctcatgat	cctcccgcc	tagcctccca	aagtgtctgg	attacaggcc	tgagccactg
80341	tgcccagccc	aaacctcttt	tttctttata	aattgtctcag	actcaggtat	tttttcatag
80401	cagtataaaa	gtggactaat	acaggcacca	tcgaattgat	tgagagccca	gatagaacaa
80461	caaggaagag	gaaagggtgaa	ttatctcctt	ctgaaactga	aatatccttc	ctccctgcc
80521	cttgacatcg	gagcttttaga	gttacaccat	tggtctccct	gattctgagt	cctttggaca
80581	tggaactgagc	catgctacca	gctttccctg	ttctccaact	tggaagacagc	ctatcgtgga
80641	acttctcagc	ctccataatt	atgtcaacca	attccccata	tgagtcttct	ctcatctatc
80701	tatctacata	tatcctattg	attctgcctt	tctggagaac	cctaattgtgg	ttacaataac
80761	acaaaattca	ctagtattata	tggaagactt	agtttttgcc	tttgcccat	tttatatttg
80821	tattataact	gtttctggaa	aatggaacaa	gtttctgtct	tcttcatatg	agggtctaaag
80881	cttttttctc	actaatattt	ttggagattt	ttaagatttt	cttttgcctt	gacatacaat
80941	cttatgaagg	ctgagaatta	aaattatttt	tctattttat	ttttcaggct	caagtgtttg
81001	ctttttaga	ttcttgagca	cgttgagagc	ctccaaggct	tggaaggggg	tgccataaag
81061	ttcagtgatt	atagggagtt	gagagactca	actgggaaag	gaaacgtcta	aacagaggca
81121	atttgagaga	taaaagtfff	ctcaaaggag	ccattaaagt	tctaaataat	tcttagtaaa
81181	gtcatgcaaa	caggaaagga	agtagacagg	attagttcct	tattggtgga	acacatagtc
81241	agcggagggt	tggaagggga	gatttttagtc	aactgagaag	ttcccatgaa	aggagcaaga
81301	tcaagatctg	atggaagggg	aagagacacc	atgaacaaa	atccaggaat	aagttccaac
81361	ccaagagag	aacagagagg	cctcaaaaac	aaagctagga	taagaaaact	ctagcctgag
81421	agttaccttc	tagacaaaga	agactgagat	tccaaccacg	cttcagagag	tactcacatt
81481	ttgatgtttac	tcaaaacttta	ggctttttta	tgacttagcc	atgcatgcaa	aaggcattcc
81541	ctaagggtggc	acagaagacg	gagcccttat	atccaaagat	agccaaggag	aaagaaagac
81601	ccctgttgcc	agagccagtg	gataaaggca	acagaaaaag	agacaagggt	ccttatgtga
81661	tgagaccttt	ctagatttag	gcttatataa	actcctgaga	actgggagga	tgcatagccac
81721	agatggggta	ccaacatttc	tactcatttg	attacaagtt	ctcgggcac	caaaatgatt
81781	aacaaaatga	caatttctag	ggcttctgtg	ggatagtatg	gaaggtcttt	tgaaacgttt
81841	taatgctgtc	aactgaagag	tgatgaagtt	catatatttg	gaaaggaaag	atttctttat
81901	ttttatgttt	ctttattctt	tttttttgag	acggagtttt	actcttggtg	cccaggctgg
81961	agtgcattgg	catgatcttg	gctcactgca	acctccacct	cccagggttc	agtgactctc
82021	ctgcctcagc	ctcccaatta	gctgagatta	tagatgcata	ccaccatgcc	cagctaattt
82081	tttgacttct	tagtagagac	agggtttcac	catgttggtc	aggctggtct	ccaactcctg
82141	ctgttagatg	atccacccac	ctcagcatcc	caaagtgtct	ggagtacagg	catgagccac
82201	tgctgccaat	gagagattta	tttctataaa	agggttacag	cctgcagggt	agttcttctg
82261	acaggtctagg	aagtatagcc	tccagccaga	agccagaaac	agacattttc	aaatgtgagt
82321	taaaggaaac	agtaatttat	gctgagtggc	atggccaaat	acacatattt	aataagctct
82381	aggaggagtc	atgaatatatt	atgaaaggag	aaatgcgtgc	atgcgcaatt	gagtgtcttg
82441	ctccttcatg	ggtcccatgt	acaaaaattg	gcagtgttag	catgatccca	ggatggagtt
82501	ttcagccccc	taacactaaa	aggtagaaga	gaggacatga	aaactcactc	tgtgcatcct
82561	ctgtacgctg	gcagaaccc	ctccgtcatg	ggtggtctct	tatcaggcaa	gaaaggagag
82621	gcggcttcag	gcagttgggt	gatatcagtg	gtggagtctt	tttcaagggc	tggtttctgt
82681	taaactcctta	gggaagaaag	cctcatcatg	gttagcaaag	gagagggtat	aacaagggtg
82741	atctgactcc	catcatccca	tgctggccaa	gctgagaact	cagttttgaa	agttactctt
82801	gggtcccttc	agccaagagt	gggtctgttc	agtcagttgg	gagcttagga	tttcattttc
82861	atttatcatt	gctaattggga	aagggtacgc	tgtctccatg	gcagctgaat	tcgcaagaaa
82921	ctccttggtg	gggtttaatg	gcagYtgat	ttttctgKga	SctYKgttt	aattggataa
82981	agtaagttct	ggtaagattt	cttccNttat	cttcagtatc	tcaaRtggtt	tcaYttaaat
83041	aatctttata	acaacttttg	atgtctgagt	ggaKtccac	acagtcactc	attgtaagac
83101	tttctgattc	cttttttttc	ctttggtcat	tMtgaatagg	gcttctgtaa	ataaYtgcat
83161	ggtagctttt	gWtKggaaat	aacatcaaag	tagttgtcaa	aatacYtagg	aatgtKattt
83221	ttggattgta	aggKgagact	tgtttagctt	tRgaaaaaaM	tgMccaactt	gtaatgggga
83281	ggaaaaaaat	tttctNtggt	tttggaattc	ttagatggRa	cScKctgtaa	aaactgacag
83341	attaaaaatga	gaaaaaSaga	aaagtttaaa	aacaYgtata	YcttatggWt	acatggggaga
83401	tactcaggga	aaaatgagta	aatctccaac	agggtggctt	caRttcaagc	ataaaacta
83461	tcttcaactt	aaagaaagaa	gatttgaggt	gcagtggtNa	gtggggagtt	aaccagcaaa
83521	agcacattag	acaagggtaa	gggtYgttat	acagacttaa	gtccaYgcat	tctccattga
83581	taagactctt	Yagtgtattta	gttatccttc	tcttcttgKt	gtcgagagag	gtagctttta
83641	aatggtgatt	tccttttatag	atgtaaaatt	tccttacaca	agtgttaactt	ctactctRtt

83701	ttcacaactt	cgtttagcatt	tttttttttt	tcaaaaataat	cagcttgga	taattcttaa
83761	gccaaaggga	catattttgK	gggKgcata	tctggtttcc	taccattata	ttttgggggtg
83821	gcataKtttg	gYcttatata	ctgtgttcca	cYggcaatga	aaagagttct	tgtttttccct
83881	ccagcaattt	gtcattttgt	agtttagcag	ttctaagagc	tatacaccag	ctgtgctatc
83941	tcactgtggg	tttcgggttct	ctagtatgtt	gagcatcttt	ttgtatgttt	acttgccatc
84001	tgtagatctt	ctttgggtgag	gttctgttca	gatctgtgtg	catttttcaat	tgggttggttt
84061	aacttattgt	ttagtttttaa	gattttttta	tatattttga	atacaaattc	tttctcagat
84121	ctgtattttg	caaataatttt	cttcaatatg	tggcttgtct	ttttgttctc	ttaacaaggt
84181	ctcttccaga	gtataaactt	taaatattaa	gaaatccaca	ttgtcatttc	ttctgtgtat
84241	atcaaccttt	tgtgtcattt	gttaaaattc	attaccaaac	ccaaaggcac	atagcttttc
84301	ttctatagtt	tcttctagaa	atggtacagt	tttgcatttt	tagtgtaagg	atgattttga
84361	gtgattattt	gtgtaagttg	ttaaagtttt	atctacacac	atatcatttc	ttatggtttc
84421	caattaactg	ttccctattt	ctgggaaaga	cacaggatag	tgggctctgt	tagagtagat
84481	agatagctag	acatgaacag	gagggggaag	ctcctggaaa	agggaaagtc	tgggaagcct
84541	cacctggagg	taccaccaaa	aattcacata	ttagtagcat	ctctagtgtc	ggagtggatg
84601	ggcacttgtc	aatttgtggg	aggtagggaga	agaggtaacct	atgcagaaag	aaacacctta
84661	gaattcctct	taagatgccc	caatcatcat	tcattctgca	ataaaaaatgt	catacatcta
84721	ctctactgca	cccagccctc	ttctgcaatt	tcaataatca	attgtgctat	ttgcctttct
84781	ttcagcaatg	agatttttatt	tttctttcct	aattatttca	aacatgaact	ttggttccag
84841	agaactagta	tttcccttcat	ttataaattg	agggcagctg	ggcatggtgg	ctcacgcctg
84901	taatcccagc	actttgggag	gccaaaggcag	gcagatcact	gaaggctcagg	ggatcaaagc
84961	cagcctggca	aacatggtga	aaccccatct	ccactaaaac	tgcaaaaaat	agccagccat
85021	gggtggcagg	gcctgtagtc	ccagctactc	aggagactga	gacaggagaa	tcgcttgaac
85081	ccgagagggtg	gagactgtgg	tgagccaagg	tcgtgccact	gcactctagc	ctgggtaaca
85141	tagggagact	ctatcctcaa	aaaaaagata	aaaaaattga	gggtcatctc	acagacgatc
85201	taataatgaa	ttattttttt	gtcttttagaa	aatcaacatt	aacttttcta	cttttagata
85261	tcgtaactgc	tgtgacttga	aggacttatc	tagaaaaagc	cttaaaaaac	tacggctcagc
85321	actgggtgaa	tgggttgggg	gaaccacat	aaaatcccca	agacacctgg	gagtccatgt
85381	ccccatgagt	gggactgcag	gcagctgtag	cagactggat	gggagagggc	agcaggcagg
85441	agaactcggg	gtctggagtc	cacggttcta	aggccaggga	aaaccactgg	caaagtgaaa
85501	tgcggagctt	gacaggatga	aatttgtgat	tgtaaatgaa	tatttgccat	ttccaagtga
85561	gatcgccagt	gggtggtggga	tggatgggtg	cccctccaag	tgggctgcag	tgaggagagc
85621	gtggcaccag	gccaggatgc	tcctgccagg	aacacaggat	ctgcacacgt	ttaggaggaa
85681	acgctgggca	gacccagctt	ggagtcatct	ctgctcttta	catctgttaa	ggctgtgaaa
85741	actgagagtc	ggccggatgc	agtggctcac	gcctgtaate	ccagcactct	gggatgccga
85801	tgcgaaatgga	tcacctgagg	tcaggagttc	aagactagcc	tggccaacat	ggtgaaaccc
85861	catctctact	aaaaatacag	aaaattagcc	gggtgtggtg	gtaggtgcct	gtaattacag
85921	ctaactcgga	agctgacaca	gaagaatatc	ttgaacatgg	gaggcagagg	ttgcaatgag
85981	cagagatggc	gccattgcac	tccagcctgg	gtcacagagg	gagactccgt	caaaaaaaaaa
86041	caaaaacaaa	acagaaaaaa	tgactctcag	gaacagttcc	cgagaaggaa	aattgggccc
86101	gcatggaaat	agacattttt	ctcccaccta	gggcaggagg	tgaagtgaag	taggtctgtg
86161	gagtggactt	tcacatagaa	accatgtatt	tcctaaattg	ggggttattt	ggggatcacc
86221	tggaggagta	ttcctgggtt	tggtgaaaca	cacgggggta	ttttttgtga	agctgcaaat
86281	ctggcacagc	aataacRgct	ggggaactgg	agatMaagga	gaaggcatac	taagtgtctg
86341	tgaagttttc	ccagaagtat	gacattattg	ggaagtaaac	tactttttta	aacacccgtg
86401	gcaataccac	gtcagtaagS	cagagacaac	accatgaagt	ttcatgacag	aggccaaatS
86461	cagcccaaac	ccagcccagc	agaggctgtc	aactcagcgc	cccagcgaga	gccggaaggt
86521	tccatcctca	gagctgcaga	ccctctcgtg	tgggctgcaa	aggccatgtc	tgcattcccg
86581	gcggatgtga	cgctctgaga	gatacatgcg	tgttccgggg	gttatatgag	tgtgacgggt
86641	gtggcgtgag	tctgactgtg	tcacgggcgt	tccagggggt	acgtgtgtgc	tctgagggac
86701	acatgcgtgt	ttcRgggRtt	atatgaggtg	gacggStgta	gcgttaggtg	acRatgtcat
86761	ctccgMgttc	caagcgttat	gtgcgcactg	agggacacat	ccacgttccc	ggggttggat
86821	gtggaaggca	gctaccccga	cggtgtgtgt	ctctgcatac	gacgggtgct	aacactagca
86881	tcacagatgc	agtgttatta	gtactacgga	ggttattatc	agtgtggcgg	gtgttcta
86941	tgttttccgt	acgctacatt	tctgttccaa	gaccgcagct	tggccctgtg	gctgctcgc
87001	cttgggtgtg	gagaatgaac	ctcagagtgcg	ctggattcac	aggggatttt	ggtttcta
87061	tttccacatg	aagggccaaag	gcgggtggat	cacttgagat	caggagtcca	agaccagctt
87121	gggcaaatag	caagaccctt	gtgtctacaa	aaaaaaaaaa	aaaaaaaaat	tagatggcat
87181	gcttgacagt	acctgtagtc	ccagctaccc	aggaggctga	agtgggagga	ttgcttgagt
87241	ccagaagttt	gaggctgcaa	tgagccagcc	atgatcgac	cactgcactc	cagcctgggc
87301	gacagagtga	gacccctgtct	ctctctctct	ctctctctct	ctctctctct	gtcacacaca
87361	cacacacaca	cggttaaattt	gttggtattat	atatttaggg	gggttgagcac	ttttgggttat
87421	aaaatattta	tgattgtggg	aacaagttaa	taaaagacacg	aaacttattt	aatgtccca
87481	gaactttaag	aacaaaaagc	attcttaggt	taaaaataag	ttttacttta	aaggtaatatg

87541	tacacacata	aattgttgtt	aaaatcgaca	gtaacaaaga	gaagtaacaa	tactaatggc
87601	ctgtcacaaa	ctgatttctta	ataacctata	taaacaaacg	ttaagcccgg	gcgctgggtg
87661	gctcatgcct	gttatcccag	cactttggga	agccgaggca	ggcagatcac	ttgaggccag
87721	gagttccaga	ccagcctgac	caacatgggtg	aagcccagtc	tctaccaaaa	atacaaaaaa
87781	attagccggg	tatatgtgca	tgacacctga	atcccagcta	cttgggaggc	tgaggcagga
87841	gaatccttga	acccaggaag	cagaggtcgc	agttagccga	gatcatgcca	ttgtgacagg
87901	agagaaactc	tgtctcaaaa	aaaaaattat	atgtttacaa	caggtgcatt	cctcctcttg
87961	ctttctgagg	acgcccctgt	atgtagctga	gtagtcgcta	ataaaactatc	ttaacttcac
88021	tatactctgt	gacttgccaa	aaggtctttc	ccatgtgaaa	tccaaaaaac	tattcttggg
88081	gtctaggaca	agacccattt	tataatgaca	aaactataca	aattctagag	gaaaacatgg
88141	aaagaaaagc	atgtgacctt	gSgtttggcc	atgagtttta	acacgacact	atcagaggca
88201	tttgaacccc	tcctctaYat	gaactccagg	gtggtttatg	ttccaYtggc	taagRgaaag
88261	ttttcttcaa	aaatgtgaca	tgatttgagg	tcaaacatta	atatcaagKW	aactcaaaag
88321	attgagaagc	tagcattagt	tctgggaaaa	ccagaagtgt	gccttttttt	ggaaataatc
88381	attggtagca	caaacttaag	aatctccaaa	ggaaataaaa	atgagttatt	aacttacagt
88441	tttcaccaat	taagatataa	atgaagctaa	Ygaaatccgg	aaatacaatt	tcactgtttt
88501	taatgttcat	taaaaaaaaa	aaaaaaaaacc	ttatcaaata	gccccagtaa	gtcaccaatt
88561	aagtctttac	tacttaaaaag	caaaaatccac	ctatgtcctg	aacacttatc	cacttttacRa
88621	gcctcattat	atgtactgga	gacaaaaatc	agaaataaat	aaatatatat	gtacatatat
88681	acaaatatat	ttcaaatttaa	aaaatacttt	tagagagtgg	tatgtattac	atttagaaat
88741	taataacgaa	gtaaattatg	ggatgtcctc	cacgcctgtc	ccaaaggtag	cgaattttata
88801	aatcatctca	ggtgaggagc	aggacagggt	gaaaatagga	atgacatgaa	ccgcgcggga
88861	acagctgccg	gcgcgggtgtc	cagggcggca	ccccgccgg	tccYggcccc	tccagccctg
88921	ggcccgacc	ctactacgcc	tctgcctcga	cgcgacgcg	gagcccgagc	gcgcgtcacg
88981	ccgtgtggg	ccgaagaggc	tgctacccag	aggcgagtg	cgggctcgcg	agggtcccca
89041	cccSactctc	gctcccgcca	gcacctacgg	actcgcgtcc	ccgcgcgcg	ccgactcggg
89101	agcagcaccc	cccccgccac	aggagcctca	cgcgccctct	acctaacagg	aagttgggtg
89161	gaagcagcgc	ggaccacagg	cacaccgaac	gcactccaac	agaacccgac	gcagacacgc
89221	gctttcaacc	ggcggagaca	ctggcagggtc	aatagagata	ttgactatat	aaacaaaaga
89281	atgacaaatt	aatagtgtaa	tgataaactt	gactttggca	aatattgtga	atttttgtga
89341	aagtacaact	aaaaggcaat	gtactccaa	taatcaccag	agtaatcaat	ttgcttattg
89401	ctgtcccttt	aaatatagtt	ctctgggtatc	aactaacatg	tttttaacta	atgatgcttc
89461	ttaaagaaaa	gggaaaagac	ctttttcttt	ctttcagttc	tcaatgattc	actgcttcac
89521	ctcgctccac	caaagataaa	tgaatctac	atctcttata	cattaacaat	gcatgacaat
89581	ttacaaatag	ctaaattttt	ggagctaaact	ttaagtacct	gaatggaatt	taatcaacc
89641	actaatctcc	ttctcacttc	tcagtatttt	atcaaagttta	tgtcaaggga	caaggaaaaa
89701	ttatccaaac	attgttttaa	acaatcatca	ttaattagta	acacttatcc	aggggggttt
89761	ttaacctttc	ccccactcaa	ggattattct	aatgtcagag	tagaataaaa	aataagtgc
89821	gcgatgctga	ctcttccaag	cttaacattt	ctcacaagtc	aattagcttt	gtactgggag
89881	gagggcgtga	agggctgctt	gcggtagtgt	tgtagcagca	gcacaatggc	cgcagacaag
89941	gaaaacagtt	tctaaaaatt	cctcgtatat	aattttatat	ttttgacaag	attaatgacc
90001	catgctccct	tcctctccat	ttcttttttt	ggaattctgt	tggtatgtag	ttactatatt
90061	ttattaaagg	aaattagcct	tatctctttt	tattttttat	ttaaagaaaat	tattatatta
90121	ttcctttata	tttttattaa	aggattttat	tattattaaa	ggaaattagc	cttatctctt
90181	attatatttt	ttatgacctt	caaagtagtg	tctctgctta	aaagtgtacc	ctggccgggg
90241	ttggtggctc	acacctgtaa	ttccagcact	ttgggaggcc	gaggcggtg	gatcacgag
90301	tcaggagatc	gagaccatcc	tggttaacac	ggtgaaaacc	cgtctgtact	aaaaatacaa
90361	aaaattagca	gggcatagtg	gcgggcgcct	gtagtcccag	ctactcagga	ggctcaggca
90421	ggagaatggc	gtgaaccggg	gagacggagc	ttgcggtgag	ctgagatcgc	accgctgcac
90481	tccagcctgg	gcgacagagc	aagactccgt	ctcaaaaaaa	aaaaaaaaaa	NgtRtaccct
90541	gaagcacaca	tcaagcgaca	tgttaggttc	ataaattctg	gccaaatggt	catacctcaa
90601	acctcatcag	caStaaggct	ctttacttgc	actgacaaat	atgaacgctg	gggaatttgg
90661	aaatgatata	taatatataa	tatttatatat	ataatagata	tataatatat	aatatatata
90721	atacatatat	aatattatat	atgtaataga	tatacaatat	ataatatata	atagatatat
90781	aatattatat	ataatagata	tataatatta	tataataatag	atatataata	tataactttc
90841	catgtgattt					

ELP3 genomic sequence (SEQ ID NO: 2)

>8:27927301-28022150

```

1      gagaaaataa ttcaaagttg tatttaaattg ttcagagggg gaataataacc aggatccttt
61     ctcaaaaaaag aaacacatgt ctggaatagg cagtgatgtg ctgcagccag ctcgtaggag
121    ctcttgaaaag ccaacagcac acatcacttt cctactctgc attcagtgac gtcacactgg
181    tagcttgaaa ttgacaatgt tgggactggt Ycagtcctga atctcaaaact gataaccctaa
241    atccataacc acattgtgat ccctaattctt caaccaataa attccagctg acatgtgacc
301    acagcttatc aatactaaga ggcacatcat ttccctgttt aagatttctg aaattgagat
361    gcatcttgca actgctatag accaggtgac aatcacaaagg tgggtgtcac tgccctgcgtg
421    ttatgggctg aaatgtatcc cctccaaaaa ttcaaagtgt gaggtcccaa tcYccactaa
481    ctctatttta caccacagga attggcaaat ggtacacatc agccttttgg gttttttgtt
541    tgtttgtttg ttttggggga tttttttgtt tttaaacaaa gagctagtgt ttaaactatt
601    acaagcgcaa cgcaacacaa acctggctaa ggtcctagaa aagtgcagt ttcttaagta
661    gttagtgagt aatcatcacac attttaatgc agactaatgt tgaacatta gtctgggtgc
721    ggtggttcac acctgtaate ctgacacttt gggaggccaa ggcaggcgga tcacaaggctc
781    aggagttcga gaccagcctg gccaatatgg tgaaacccca tctctactaa aaaatacaaa
841    aattagccag gcatgggtgt ggcgacttgt agtcgcagct actcgggagg ctgagacaa
901    agaattgttt gaacccggga ggtagagggt gcaatgagcc aagatcatgc cactgcactc
961    cagcctgggt gacagagaga ttccatctca aaaaaaaaaa aaaaaaaaaa tcaagcaatt
1021   aagaattatt tgagggagtc tttttttgtg cttagaaaaa cagtgtcgtg gagaagtcatt
1081   actagaatca tatagaaata gcattacata ttctttcatt cttcttactt aaaaaaaaaa
1141   agaaaaacat tcacctatag aatgtgcatg tgctgtgtgt ggtatttgag ggaggtagaa
1201   agcagtgaag agttgtcaat ctctttctga tgcaaagata taaagagctg agtgagtaga
1261   cgccttgca g tccactggga ctggtgagaa gccgtgtccc tggcaagcta taagcacgtg
1321   tagttaaggc tggaatacgg atatctgtgc agactgccag aagcccaccc ccaccaaccc
1381   ccttagtaga tgagctgaag ttgcaaaaag aatctatctg ctgagatttt ctgttgattg
1441   agttttcact ctaggaaaaa gtttagggaa agctgaaact attttctact gggagatgaa
1501   attaatgtag aattaagttt aagtagagtt ttaaaRcagt tccaagaaag gcatgtgctt
1561   agctgtgtcc agatgtttgc attatttaaa atctggacct aaatcttgta taacttttct
1621   aaaaattact gaagattagt taacatttat gctttttatt gtagtaaaaa gaaatgtttt
1681   tatagtgtca gtgaagtga c tctagataaa agggttcata ttattgtaaa agagaagtag
1741   aaacagctac agctagaggt aggatccaag atgttagtag atattcatta aattatgcaa
1801   cagaaacatt acacactgtt ctaatttaaa cacatggctt gtcataataa aacgtaacac
1861   aatatattaa agtttgtaaa aagatgttgt atgtgagaga ggagaaaacg tggagactca
1921   aaccagaagg gaggttaggt gtggtggctc acacctgtaa tcccagcact ttgggaagcc
1981   gaggcagggt gattgcctaa gctcagggaat tcaggaccag cctagccaac atggtgaaac
2041   cccgtctcta ctaaaaaatac aaaaaattag tcgggcatgg aggcacacgc ctgtagtccc
2101   aactactcgg gaggtggaag caggagaatt ccttgagcct ggggaagcaga ggctgcagtg
2161   agccatgatc acgccattga actccagcca ggacaacaaa ggaagacact atcttaaaaa
2221   aattaaataa aaggccgggc acggtggctc atgctgttaa tcccagcact tcgggagggc
2281   gaggcgggca aatcacggga ggtctggagt ttgagaccag cctgatcaac atggtgaaac
2341   cccgtctcta ctaaaaaatac aaaattagcc aggcattggt gtgagtgcct gtaatcccag
2401   ctactcggga ggctgaggca ggagaatcgc ttgaaccagg gaggcagata ttgtggtaag
2461   ccaagatcgt gccattacat tccagcctgg gcaacaagaa caaaactctg ttaaaaaaaa
2521   agaaaaaagg aaaggaaaag aaaggaaaag aaaggaaaag aagaaagaaa acagaagggt
2581   tattcaatgg aagcagaatc agaaaaggata tagtaaagtt ggagttacac tgaaccagt
2641   aatatcaatc atgacccagc tgggtgaggt ggctcatgcc tghtaatccta gtactttggg
2701   aggctgagac aggtgggtca cctgaggtea ggagttcgag accagcctgg ccaatatggg
2761   gaaaccctgt ctactactaa actacaaaaa ttagccagtt gtggtggcgc atgcctgtaa
2821   tccagctac ctgggagact gaggcaggag aatcacttga acccaggagg cagaggttgc
2881   agtgagccaa cattgtgcca ctgacttcta gcctgggcaa cagagccagg ctccgtctca
2941   aaaaaaaaaa aaaatcatga cctgattttt aaaaatctac ctgaaatcag tgacataaat
3001   catgactcat tttttaaatc caccttttct agctttgtca tgaaaattat gttgcagtga
3061   tgtcactctg ccactgtctt acttgtatga aataatttga tctgagcccc tagcactctt
3121   tttttgtctg tgttagagat ggggtgtctg ctatgttgcc ctgactgtgc ttgaactctc
3181   ggccctacac gatcctccca cctcggcctc ccaaagtgtc gagattacag gtgtgagcca
3241   ctgctcctgg cccaattatt atttttaatg ttattctcgc taaaagaaac ccagattcc
3301   tggggagtag ttgttccata tcttggggca gggaaaaata aagtgaacct gaatatcttg
3361   ttgtagctag acagcaagaa tgttttttaa atgctaaagg gagtaatgct aaaggagtc
3421   agcttaagg gttccactg gccaaaaagt gacacattga gcatcagaat gaacacata
3481   gttctgtgaa aagctatgag ctcaataatga tcttaaaact gagtaataa aataaggat
3541   gttaaagtaa aggcattaac cacctcaaac aatatatgcc tgagactgtc aaaattcagt

```

3601	ctgtaaatttc	cacttgggca	agtcctaaaa	gtttacctaa	catataaaaag	aatattaaca
3661	gtcttagtag	gggcacaaaa	atcctgtcag	gtcaaatcag	gcagtaaaac	tttgtgaatt
3721	agggttaaaat	aaaaagtcgt	gtgacagttt	atagtaaaaa	ctgagccaag	gcgggccaggt
3781	gcattggctc	atgcctgtaa	tcccagcact	ttgagagcct	gagggcagat	cacctgaggt
3841	cgggagtttg	agaccagcct	gaccagcatg	gagaaacccc	gtctctacta	aaagtacaaa
3901	attagctggg	cctggtggcc	catgcctgta	atcccagcta	cttgggaggc	tgaggcagga
3961	gaattgcttg	aacccaggag	gcggagggtg	cggtgagcca	agatagcgcc	attgcactcc
4021	agcctgggca	acaggagcga	aacaaaacaa	aacaaaacaa	aacaaaacaa	aaaacaaaca
4081	aaaaaaacca	aactgggtca	aggcaaaatt	taagggaagg	tcaagtgaac	cataacatac
4141	caggtagtaa	gaaaagatgt	ggagctagaa	agcctcgtta	ctataggatg	gggcccaaat
4201	caggcatcct	ggaaggatct	atgtgcctgt	gtctcagcct	tgaatctcaa	actagtaact
4261	caaatccata	actacattgt	gatccccaat	cttcaaccac	ttaacttcag	ctgatatgtg
4321	actacagctt	atcaatacta	agaggcacat	tgtttccccc	tttaatat	ttgaaattga
4381	ggtgcatcct	gcaactgctg	caggccaggt	ggcagtcaca	agggtggtgt	cactgcctgt
4441	gtgttatggg	ctgaaatgta	tccccctcca	aaattcacat	gttgaggtcc	taatctccac
4501	tacctaaaaa	tgtgacctta	tatggaaatt	gggtcattga	ggatgaaatg	tattaagatg
4561	aagtcataat	ggagtagggg	gggcccctaa	tccaatgact	ggtgtcttta	aaaaaagggg
4621	aaatttggtg	acagagacac	acacatgagg	agaatggcat	atgaagactg	gagttgtgct
4681	gccacaagcc	aggaattacc	agaagcaagc	agacagaact	gtaccagatc	cttcctgcag
4741	tttcagaggg	agcatggcgc	tgctaatacc	ttgatttttg	acttccagcc	tccagaactg
4801	ggagataata	attttctgtc	gttccaagct	acccagctgt	ggtgcttagg	aggttctatac
4861	actgcacatg	ggtgaacttg	gtcacagctg	ttcacattgc	ccttccttct	aatgagttgt
4921	gtgcattact	taactgccat	ttaatatgtc	ttttaaaaga	ttatgttatg	attcagcatt
4981	gaaaggctcg	cattttccag	gccgttacat	cgaaggccct	ggaaatggaa	acgtggggca
5041	tatacttgat	gttagtggaa	gtatggttgg	tgtagacagt	aattccagat	ttttttgcaa
5101	agcaacacct	aagtgcctca	tgggacctag	gaaagggaga	catccacatg	tagatgaagg
5161	tggtgtgtgt	tttgatactg	aggtgtatgc	aaagtgactg	cctatcccac	accagcccaa
5221	aaaaaaaaagt	gatggaaact	caaagaatgc	cttataaaga	acttctgtct	catgttttct
5281	gtgattaaaa	ctaaattttg	tccaaggcta	ggcattgtgg	ctcacacctg	taatcccagt
5341	gctttgggag	gcccagcgag	gagaatttgt	tgagcccagg	aattagagac	gcagctgggc
5401	acataacctt	gtctctacaa	aaaattttaa	aattagctgg	atgtggtggt	cgcataccgt
5461	ataccggtag	tcctagctcc	ttgggaggct	ggtggtggtc	agggggtgcg	gagtgcgggc
5521	aggaagatag	cttgagccca	ggaattcaac	gttgacgtga	gttatcattg	tgccactaca
5581	ctccagcctg	gctaacagag	tgagacccat	ctgggaaaaa	aaaaaagttc	caaacaacRg
5641	catgtagtac	aagtaaaggt	ttcacaataa	ttcaaactga	aataaagtg	ttggagaatg
5701	aaggcatcac	cttgccatca	gaggtgtggg	ttttaggcct	cagaagaact	tgaccatag
5761	cggtgacagc	tatagtcacc	ctgcaaaactg	atgtattgct	gcccaggccc	taactcatac
5821	aaaccaattc	ggcatgttta	ctggcatagt	tgtaaatgtg	gtattgaagt	attctacgtt
5881	tccattttcca	gggataatat	gatcaggcac	tcatacaatg	catatttatt	taataccaac
5941	aagaacaatt	caacagacca	agaccagctg	tactaagttt	tgccctgcta	aaactggcat
6001	tcaagaattg	cctgatagaa	tattttggct	taagtatgag	gcctaagtat	attgcgctag
6061	aaacaaaagc	tttgaaattt	tcaattccac	gttgacacac	agacattttc	tgtatcaata
6121	ttattgaagg	caaaatacag	aaagaaattg	aatatagagc	tggtcatcgg	ccaaatgttt
6181	ttactattga	atccaatacc	gaagattttg	tttcagccaa	acataatcatc	tgtcacatgg
6241	tacttacttg	gatataatcc	ttatgttgtg	ttgtttatat	tttatcttaa	acttggtttt
6301	agttttatag	ttgtacaaga	ctattatagg	ccagacgtgg	tggtctacgt	tggtaatccc
6361	agcacttttg	gaggccgagg	tggttagatc	accagaggtc	aggagttcga	gaccagcctg
6421	gctaacatgg	tgaaacctca	tctccactaa	aaatacaaaa	attagctggg	catagtggca
6481	ggcacttgta	atcccagcta	cttgggaggc	tgaggcagga	aaatcgtttg	aacccagggg
6541	tcagaggttg	cagtgcagcc	agattgcacc	actgcactcc	agcctgggca	acggagcgca
6601	actctgtctc	aaaataaaaa	agttaaaaaa	aaagactata	ataataataa	gttgtatgtt
6661	tgtacctact	ttagtataaa	aatagtaaaa	ataggagaa	actaagtga	gctcaggcct
6721	tatccctgtc	cactccctct	cacactctcc	ctctgagcct	ttccaagcta	cttgccctgtc
6781	attcgccaaa	cacagcagcc	tgtgtgaacg	gtgttctttg	ttctgectca	aaagtgcctt
6841	ctcgactttt	tcaccttact	aaagactaca	gcttctttaa	aattctgtct	tgagagcac
6901	tgtggtgctt	gccgagctgt	agtcacagct	acgtgggagg	ctgaggcagg	aggatcactt
6961	cagcccaaga	gtttgagacc	agcctgggca	acatagtgga	aaccccatct	ttaaaaataa
7021	taataataaa	attggccagg	tgagtggtgt	catgcctgta	atcccagcat	tttgggaggc
7081	caaggcaggc	agatcacctg	aggtcaggag	ttcaagacca	gcctagtcaa	catggtaaaa
7141	ccctgtctct	actaaaaata	caaaaaataa	ccaggcatgg	tgaccagcct	ggtcaacatg
7201	gtaaaacctt	ctctctacta	aaaatacaaa	aattaaaccag	gcatggtggc	atgcacctgt
7261	aatcccagct	actcaggagg	ctgaggcagg	ataatcactt	gaacccagga	gacggagggt
7321	gcagtgcagc	gagtttgtcc	cactgcactc	cagcctgggt	gacaaaagtg	gactaagtct
7381	caaaaaataa	taataataaa	ataaaattct	gctcagaggc	tgagttttcc	ttgaacacct

7441	ccccagtgccc	acctctcctg	ccaccagtg	gacctggcat	ctttcctccg	tgtcccatag
7501	cactcaacaa	ttcctccact	gtatggagag	cccagactca	cctgggccta	ctcctctgtg
7561	gaccacatgc	tcattcgacc	ctctcaacaa	tcacgggaga	tagatatcat	tatctccctg
7621	ctttaaagat	gagaaaatgg	tgcctcagag	aggtaactga	agcagtagga	ctgaggtaaa
7681	tccaaaactgt	ctcctatagg	tcagtaatcc	cagcactttg	ggagggtcaag	gcagggtgat
7741	cacttgagcc	caggagtttg	agaacagtct	ggacaacatg	gtgcaatccc	atctctacaa
7801	aaaatacaaa	aattagccag	acatggcggg	gcatgccctat	agtcccagct	acttgagagg
7861	ctgagggtggg	aggatctctt	gagcctggga	ggcagagggt	gcagtgagtc	aagatcacac
7921	cgctgcactt	cagcctggcc	aacagagtga	gacccgcgtc	tcaaaacaaa	agaaactaac
7981	tgtctattcc	aaaacttttg	cccttgcaac	tacacaatac	caggcataca	tttgtctccc
8041	aaaatctagt	tgctgggata	cagaaaaaca	catagagaaa	cctattaact	caccttcagt
8101	ttttccttga	gaattctagt	cccttgtagt	tttatcattt	catttctttc	taaaacagcc
8161	tctcgctgac	tctgaatagc	ttgctagggt	gtgaaaggaa	atagaagcat	gcagagagta
8221	gcgttatcaa	cagtgatgag	gtggcaagat	gcacattttt	ttggcagatt	ttcattttgg
8281	ccaaacagtc	aaatatgcat	gcactgtatt	agaattatac	acagttgccc	ggcacgggtg
8341	ctcatgcctg	taatcctagc	ccgttgggag	gccaaaggtag	gcagatcaca	agctcaagag
8401	atcgagacca	tcctggctaa	tccagtgaag	ccccatctct	actaagaata	caaaaaatta
8461	gctaggcgctg	gcggcacgtg	cctgtagtcc	cagctactcg	ggaggctgag	gcaggagaat
8521	tgcttgaaca	caggaagcag	agcttgacag	gagtcgagat	cgcgccactg	cactccagcc
8581	tggggcgacag	agtggagactg	tatctcaaaa	aaaaaaaaaa	aatacacagt	tgacccttgg
8641	actcatggg	tttgaactgc	gtaggctccac	ttatacgtgg	attctttttc	aataattgaa
8701	aaagatttga	agatttgcaa	caatctgaaa	aaactcacag	acaaacccctg	tagcctagaa
8761	gtatcaaaaa	attaacaaaa	agggtgcatg	aatacataaa	atagatgtag	atcccagttt
8821	atatttgata	atgtactacc	acgaaatgca	cacgattcta	tcatgaaaag	ttaaaattta
8881	tcaaaaacttc	tgtaaacatc	tatagactac	atggcaccat	tggcagtcac	gaggaatgta
8941	aacaacatt	aagatgcagt	attaaactcat	aactgcataa	aattaactgt	agtacatact
9001	atactactgt	aataaacttca	gagccacctc	ctgttgctgg	tgtgggtgagc	tcaagtgttg
9061	caaatatcca	cttaaagcgc	caaacaccat	gtgatactaa	tgatctccat	gtgaacaact
9121	gctccagtaa	attgcgatg	gcaataaaaa	gtgagctctt	gcagttctca	tgtatttttc
9181	agcctgttta	gagcaatacc	ctaaatcttg	aataaacacca	tcggaccat	gtgaagtcca
9241	gtgatccctg	aggggctccc	aagtgcaga	gaaaagtcac	gacattacaa	gacaaaagttg
9301	aatcgcttca	tatgtgccgt	agattgaggt	ctgcagctgc	tgttgccctgt	catttcagat
9361	gtacgattca	tcttgtaaac	agatgaagta	aacttatgct	attgataaat	acgggtgccgt
9421	attgtaaatg	tattttctct	tccttatgat	tttctttcct	tttttttttt	tttttttgtt
9481	gttgttgttt	gagatggagt	ctcactgttg	ctcaggctgg	agtgcagtg	cgcgatctcg
9541	gctcagtgca	acctccgcct	cccaggttca	agtgcactgc	ctgcttcagc	ctactagtga
9601	gctgggatta	cagggtccatt	ccaccaggcc	cagctaattt	ttgcgttttt	agtagaaacg
9661	gggtttcacc	atgttggttg	ggctggctct	aactcctgac	ctcgtgatcc	gcccacctca
9721	gccttccaaa	gtgctgggat	tacaggcatg	agtcaccaag	tcctgccatg	attttcttaa
9781	tgacattttc	ttttctctag	cttacttatt	ttaagaatac	agaacataat	acatacacia
9841	aatatgtatt	aatcgactgt	ttatgtctatc	ggtaagactg	ccaatcaaca	gtggctatta
9901	gtagttaagt	ttttggggaa	tcaaaaatta	tgtgtggatt	ttcaattgag	cagggagttg
9961	gcacactaat	ctctacattg	ttcaagagcc	aactgtaata	tcagtcataa	aaagtattat
10021	ttaaaaagaa	ttagttcagg	ccggggcgcg	tggctcacgc	ctgtaatccc	agcacttttg
10081	gaggctgaag	ccggtggatc	acctgaggtc	aggggttcaa	aaccagcctg	gccaacgtgg
10141	tgaaccccca	tctctactaa	aaatacaaaa	attagccagg	catgatggtg	ggtacctgta
10201	atcccagcta	ctctggaggc	tgaattagga	gaattgcttg	aacctgggag	gcagagggtg
10261	ccgtgagccg	agatcgcgcc	aatgcactcc	agcctgggtg	acagagcgag	actctgtctc
10321	aaaaaaataa	aaaataaata	gataataaatt	taaaaataaa	aataaaaaaa	actagttcat
10381	acttacattt	cctgcctttc	tttccttgaa	atgaaattag	aaaaatgtaa	caattactta
10441	atttttacta	tgtgtcactg	cagtctttcg	actttcccca	aaagcagcac	acactccctg
10501	tgtctgtgcc	ttttcataag	ttctttcttt	gacccaatgc	tctctcttct	gacccatctc
10561	cagcatccac	ttgaaaaaag	tgccttcatt	tctcaacggc	cagccagcgt	tttgtcagct
10621	gtgaacattt	tctgtgacct	cactccctgt	cccagtcaaa	aatgcagccc	tctcacctgt
10681	gggtcccata	gcacctggct	tgctccctca	tgtataactt	ttaaccctct	tttagatggg
10741	gagctcaaag	actgggcaat	gtgcctctga	atccacagcc	cctacctatg	ttcaataaat
10801	gcttatcatt	aataaattag	atttaattat	gttaattttt	tttttttttg	agacggagtc
10861	tcactctggt	gcccaggctg	gagtgcagtg	gtacgatctc	agctcactgc	aacctccaca
10921	ttccagggtc	aagcgattcc	cctgcctcag	cctcctgagt	agctgggact	acaggtgtgc
10981	atcaccacgc	ccggccaatt	tattttgtat	tttttagtac	agatgggttt	caccatgttg
11041	gccaggatgg	tctcaatctc	ctgaccttgt	gatctgccc	cctcagcctc	ccaaagtgtc
11101	gggattacag	gtgtgagcca	ccacactcag	cctaattgtg	tttaattttg	tgttggcatt
11161	ggaaagggag	ggctggagat	attctctgct	ggggtgagat	ggggagaaat	gcaggaggaa
11221	aaaaatgaag	ctgcaattcc	aacccagacc	atgacagcaa	gggaattcct	ggagacagca

11281	gtcagaccta	agcttcagtg	ggtgtggagg	aaatggtgga	ggggctgtaa	aggagctctt
11341	gttggagggtg	agagcgcaccc	gtgacatcac	atcccggtgag	cacacaccccc	atggatccag
11401	gaaagcccag	tttatacccg	ctgatccct	ctaaggacta	atcatacccc	ctctcactct
11461	caaaattgtc	ccaatataga	aggttagtta	tagtcttct	aactataggg	acaacagagc
11521	tctccataga	aatcaacatc	tattaagccc	aaaaatgcat	acctttcatt	gtaatttgtt
11581	ataatacaac	tcttagagaa	tgcattctacc	ttctcccgaa	gttctaaaaac	ttttgaggac
11641	ctgttccaag	tgagttaaaa	agttcaggag	tctctggctg	gggatgctgg	gcttgatacc
11701	taggtgatgg	gttgatctgt	gcattctaacc	accatggcac	atgtttacct	gtgtaacaaa
11761	cctgcacatc	ctgcacatgt	accctggaac	ttaaaaataa	aataaaaaatt	taatgatttt
11821	taaaaagctc	aggagtacaa	aaacctgaaa	catcaaagct	gtaccctggg	ctgggtgcag
11881	tggtcacac	ctgtaatccc	agcactttgg	gaggccaagg	cgggcagatc	acctgaggtc
11941	cggagttcaa	gaccaccccc	gccaacatgg	cgaaccccca	tctctactaa	aatacaaaaa
12001	ttagctgggt	ggaccggccg	gcgcctgtaa	tcccagccac	tcgggaggct	gaggcaggag
12061	aattgcttga	acccaggaga	cggaggttgc	agtgcgtga	gatcgtgcca	ttgcactcca
12121	gcctgggctg	cacagcaaga	ctctgtgcca	aaaaaaaaaa	agctgtacct	taccagatag
12181	gattataggc	tgtaccccct	agttatgggg	gcattctacac	tttctacatc	tgtaggagc
12241	gtgggatcat	cacatcacac	agtgaggag	gtacaccatc	tgcattctgt	acattgtagt
12301	ggcaccacca	gagctgtgta	ggggacacca	gcttggcact	aagcagctac	cctggaaggt
12361	tcagattcct	aagacaaccc	aaccaagggg	cctgatgtcc	caggaagtca	cccagccttt
12421	tctttccttc	ccaacaatca	cttcctgaag	ccctctaacc	acgttacaaa	gtcaagctcc
12481	atgtactgt	agtcaggacc	cacagaggaa	agccccctgac	tgtagccac	atcctlgcaa
12541	gcccagcaca	ataccttagg	tattctggca	agtggagttt	gtccattctg	gtgaccacca
12601	gggccacgtc	cctacagaag	ccccgctggc	aggctttgat	gctctcattc	agaaggtctt
12661	cgtgggcttg	ccccccagaa	actcgtctta	tgtcgtgat	caccagatc	actgagcact
12721	tgtcaatggt	ctgcaaaaaca	tgaccacggc	atgtgtcagc	acaggaatgg	agtcacagtt
12781	cgtgaggacc	ggttccctaa	tgtatccctt	aagcctctcc	ctccagctcc	acccctcacc
12841	gctccccacc	gcactctgtg	cactagctgg	actgagctat	gagtcatttc	tcatatgtgc
12901	cttgtttctg	taattccctg	gccacagcct	ctgcctagcc	cccatttgtt	cctcaagact
12961	gtggaggacc	tctcgggcat	tgggtctggc	tctgcattct	gtaacacctg	gacttacctc
13021	aagcagcgcc	tctgagacac	agtgcattga	ccaggtctgc	tctctgcccc	ttgatgggaa
13081	gcgactgtct	gcacctttcc	tcttgggata	cacctcacct	caagcccagt	gcttggcaca
13141	caggagaggg	ttcacagaag	tttgtggcat	tagagattga	cagaattccc	taaatcagtc
13201	ctccagtccc	agtcccagat	ccatcactgg	cttgctttat	ctcagttgtc	ccctggggaa
13261	aattcttgtt	tttccagaac	taggtttggt	ggtgcacctt	ccctcaagct	gtcacctgac
13321	ggcagcattt	tgatgtttga	ctgcccctgc	acgtttagtt	caatgccacc	agatggcagc
13381	atggagccag	aaaccatatg	acaagttgca	aaggccaatg	aacatgaaac	aacataatgg
13441	catgaatggt	gctttattaa	atatgcaact	ttcacatgac	gatcatgtgt	agccttgtga
13501	catttctcta	attgtattgt	tactctgtc	ttacctccac	ataaatgaaa	atctggttca
13561	actcagggat	tgagtttatt	tgtggttctt	cagtgccttg	catacaagg	gatgacaagt
13621	ttgtgacttg	aaaattaaat	catcattttg	ctttatacca	gtagttttta	actagtagag
13681	gttttgtgct	ccagaaaata	taaacagctc	ccaatttgtg	atgattcagc	ttacaatttt
13741	tcttttcgac	tttacaatgg	gtttactgaa	caatttgtaa	gtccagtaca	atcggggcta
13801	atgatggtta	tgattccagc	ttacagtggg	cttatcagga	tgcaattcta	tggtacatca
13861	atgagcatct	gtatttggca	aaaaatgtct	ggagacattt	tttaattgtga	caaccttggg
13921	ggggtacttg	aaggtaaagg	ccgaggatgc	agctaaacat	tcttcaatgc	ccaggacagc
13981	ccaaccccag	ctcccatccc	acaaataatt	atctgacagc	ccaaaatgtc	aagagtttga
14041	gaaaccttga	tccagggcag	tgggtttcag	cttgctttta	gcatgggagg	ctttttcaaa
14101	tgaaatctta	gttagaattc	caacatataa	acttgatttt	taaaaagcag	aactgctctg
14161	gttgaatttg	gagtgcagga	cccagaacca	ccatcggtct	aatgccctct	accctctgtc
14221	caggaacctt	ttggctccat	acatcactgc	ttaaaaagca	tctgcttctc	aagttagaga
14281	aaccccttgg	attataggga	agtaattgaag	cccaagcggt	aaagagactc	aagtagggcc
14341	acacgacgat	ggatcctgtc	tctcctctat	aagtgaatgg	actttgattt	aggcaagaaa
14401	ataaaaaacca	ttacagccct	tctgctgccc	aaactataat	caacttagaa	aagctttggg
14461	ctctggggcca	gatgtggttg	ctcacgcctg	taatcccaac	attttgggag	gctaaagcag
14521	gcagatcact	tgagctcagg	agttcaccag	cctgggcaat	atggtaaaat	ctcatctcta
14581	caaaaaaaaa	aaaaaattag	ccaagtgtgg	ttgtgtgcac	ctgtagcccc	agctactcgg
14641	gaggctggga	cagcaggatt	gctcaagccc	aggaggtcga	ggttgacagtc	agccatgatt
14701	atgccactgt	actccagtc	aggcaacaga	gcaagatgct	gtagcaaaaca	aaacaaaaaca
14761	aaagtagaag	cagcttccag	ctctgtagt	tcaacagtgg	gcagagaatg	gtccagccag
14821	cggaaggcca	taaggatgcc	aggctgggtg	ctccttaagg	gcagggcccc	cgctaccttc
14881	atattcccca	tggctcctgg	aagccacagg	tgcttagtgt	ttactaacat	gagaaaaagga
14941	cagcaacccat	aaagggtgca	gtggttaaat	ttcagaccga	ccccacacaa	agctccact
15001	gccagggttt	ggttaaagtta	gtgtcgcaga	agatacaaaa	aagcatttctg	ataaaacctg
15061	aactgccaga	cttaggagta	gatgaggtgg	cacccccacca	cccccaacct	gcattctccc

15121	aggcccccgc	tatgagaatg	ggacttagag	gagttcagag	acatgcctca	agtcacgtga
15181	cttctaagt	gcagatcaga	ttcagggaca	tctgaattca	taacgggttt	tagggaaggg
15241	aagactggca	gcctgagaga	catagggagg	gcaactgtgt	cgttctccct	gcagaggggg
15301	gtcttgagta	acttcttggt	agaggagggtc	aagaaaaccc	acaagaatgc	aggggagctg
15361	ggccaggcag	agcacgggtg	gaacccatccc	tcgcttcagc	ctgacctcct	cagagcacat
15421	tttctgagcc	agatctgccg	tctttgataa	gctgcccagt	gtatctgaac	tggtgagcgc
15481	cgcggtctcca	tccagcgaga	ggccctggca	gccttctctg	gctgtcgctc	tctcgccgg
15541	gcctggaact	agccaagatt	cgcgatttcc	accccaggat	ccgaggtctt	gcagaaacta
15601	aggctcacaa	ggccaggcct	gagccaccga	ccaaggggaa	ggacacagag	actgccttcc
15661	tcctgcagga	ggtggataat	atagagcgtg	ccctggggga	ggtggggtag	atagagcccc
15721	agcagggaag	ttaggaaacc	tgctctccaa	ggctgtgtgg	tcctggacaa	gtgactttcc
15781	ctctctgagc	cttcatgtct	tcctctgtac	aatcagggca	gaggctagag	cttgctctag
15841	attgtgactg	gaagtcgtgg	ttaccaagct	cccctggcac	aatcattcat	ctaaacaaac
15901	ctgtggagtt	taaagaactg	gattctttgg	cagctgctgt	gctgctcaca	actcactgat
15961	gctggtcccc	ctctcccagt	ccccttccag	ctcctgggct	cctgtgttcc	tcccaccctg
16021	tgctccagag	cgccctctgtt	taccgcgcc	tttctcaatc	acatgcccc	gtccctctct
16081	tcctgcaagt	tttcccagga	agccaaactc	tggagcagta	gtactaaaga	ccccagaact
16141	tttccaggca	tgtcttccat	agtctccctc	ctccacggac	cccatgctcc	ttctccatt
16201	cacccaggaa	gattcaccct	gccccacca	cgccagcttc	ctcagacctc	cagctctatg
16261	ctgatgctgc	ttacagccaa	cacagttgca	gcctcttggt	tcttggtgac	tggtgagcgc
16321	ttctctgctt	tgccccaaac	acttggtcctc	ccagttccct	tctgtccttg	gtcactctgt
16381	ttcctcttaa	gccactctat	tgctggtctt	tctcatacat	cacacacact	tgtcccagct
16441	ccaatagtca	catcctgtga	ttcctccctc	cccactttct	ctggctaccc	acaccctgcc
16501	agggccttga	gacttagaac	atagaaatta	aaagaattta	attccaatga	actcaaaatc
16561	cctgaagctg	aaaccagagc	attctttggt	gcttgagaaa	gggtagagga	gtttagggaa
16621	gagagaacta	aacctacgaa	agacagtcac	ctttttccac	atctcgtccc	tctgtctggt
16681	gaagtgcctc	gtgcctggga	tgtccaccag	cacgaccctc	tctgggatca	ggtcggtatt
16741	gggaagtgtc	acttccacat	gtttgatcaa	gggccagatg	cgcatctcag	cggcctctcc
16801	atcccaatct	ctcctctgtg	tgcggtatga	ggggtccagc	ttgatggaca	gctcttctgc
16861	ctgaagaagg	aggacagagt	cacacacacc	agccatgcac	ccagctctgt	tcaccccggt
16921	tggctgtgat	ggaggttcat	gctaagtatg	ctgagtaagg	gtaagaagtg	cgccaacag
16981	gaggcctgca	atcaggtctg	taggacctag	aacctgagag	ccaggaagga	ccctggagag
17041	gagcaagtcc	aagatccctt	ccttacagat	gaaacaaaac	caatgtcagc	taaaatggga
17101	cttttagggaa	gttatagtat	tgctgaataa	taagtttatc	ttctaaactg	cattcagttc
17161	taaaattaac	atactatgta	attgaYattt	tcattaccag	aaagctagca	atgactatag
17221	aaaattaaca	gtgatattgt	aaccatattt	tccaaactga	aaattgatct	gaaattggat
17281	cttttttttt	tttttttttt	tttttttgag	acagagtctt	gctctgttgc	ccagRctgga
17341	gtgcagtggg	gcgatctcgg	ctcactacaa	gctccgcctc	ccgggttcac	gccattctcc
17401	tgcttgactc	cccaggagct	gggactacag	gcacccgcaa	ccacgcccgg	ctaatttttg
17461	ttgtattttt	agtagagaca	gggtttcacc	atgtagacca	ggatgggggc	gatctcctga
17521	cgttgtgatc	cgccctcctc	agcttcccaa	agtgcgtgga	ttacaggcgt	gagccaccgc
17581	gcctggccga	aattggatct	taaaaagagt	ttttaaaatt	tatcttctga	ttacaaaaaa
17641	taccttgata	tgtaggcatt	tctagaatat	tttaatgatg	gtgggtgaga	gaggggaatg
17701	ttcttagacc	acggttccta	cttaaacatg	tctcgggaca	atgacacatc	catgccagca
17761	actatgacag	taaatcgcaa	ctggggtaact	tcagaggaca	ggggatttgg	gatgtattgt
17821	ggtcactggt	aaataaattt	ctataaaata	gtcctgaaag	aaagataaga	atgtcagttc
17881	ttaagagagt	ttaaattaaa	tcgctggatt	tgttgcaatc	tggggaagga	gtaagagatt
17941	tcatcaattt	caaccaaatt	ataaaccacg	acggaggaga	taataaaaagc	tctttaagac
18001	attgtaattt	taaaatagag	cagataccgt	aagagagaag	cccaggaaaa	ggcaggggct
18061	tgaaaagccc	cttcctacaa	gagggctcag	cccaataaac	tggttccggg	caaactcgt
18121	tgtagaaggc	cttgatttcta	ggatctccag	tggaaaccct	gaatttgtaa	gcccactcat
18181	cgagatggta	attctgttca	ccttgtttcc	gtgcgtaaat	ttatgacgct	gggaaggagt
18241	atttctaaat	atagagttgg	caaagacggt	gatacccttg	tgtaagttgc	gtgtattcag
18301	ataaagcaac	tagttttgta	atcattgtgt	ttcttagaca	tgctcagctc	atgttttaga
18361	tattagaagt	gtaatcattt	taatagtaat	attcgtctta	aaactgtaag	aaatgggcac
18421	tgtatttgtt	tgattttcca	ggtatgcaag	agatagtaac	atgctttgaa	aaggtttcat
18481	tcatttgatt	tttaagtgtg	cctagtcaag	catttaaaca	ttttaagagg	ctcaagtata
18541	ttgaacttta	actgtaatgc	aaaaagccta	aggaaatgtg	attaattatg	taagtatcat
18601	taatatttta	ggaagatttg	ctttgttaat	caggagatca	ttgaacattt	aacttttagga
18661	catgtataaa	aattgtaccc	attttaccaa	taatcactag	atttataaat	agaatagcaa
18721	tatttttaag	gtgtattttta	aggtgagaga	tacaagaaaa	actaggtttt	taaatgtata
18781	gcttttctaa	attgaatgtt	gattccacaa	atattttgtt	tttaattcaga	aattgaattt
18841	taacttaaaa	taaacattgt	ttataaaaagc	aagaatatca	ttctactagt	tgccatcga
18901	atgtccatcc	tgtccctctt	ttatgccagc	agaatcctga	ttttctgtct	gactggggag

18961	gcagtggagt	gaaactggtc	taagcaaato	atgacaagtc	tgcttttcac	tttctctgga
19021	ctccctttca	gccaggaagg	gctgtgtggc	tcagtcttga	actctgaaat	ctacacagaa
19081	atttgtcaga	ggtattttctg	ggaaagtttt	tactttctga	tgaaagggga	gatgtggctg
19141	atgccaacct	tttctccact	tcttccttcc	aggaggctag	catgaagtct	agcactccgg
19201	gagccatctc	atgacaatgg	ggcgacgga	aggcaaagg	agccagagat	gctgatgttg
19261	gcattgtctga	tccactaaac	caacccacc	cctactgccc	ccaaacctct	tacatgagaa
19321	aaggaataaa	tccctgtttg	tttaagccac	tataagtctt	cttaaaaaat	ctttggcagc
19381	ccaaactatc	cctaattatt	ctcaagcaac	tgtaaatagt	ctttttctgg	gcataagagc
19441	tcctccattc	acataccaaa	aatttccatt	tgatagccat	aaataatgag	gataacttgg
19501	gctaaggata	cttctaattt	gggtagcaac	ttatttctta	tttcacacaa	acactgtacc
19561	acaagcccat	cccctgaaac	aggttcaact	gttagacttg	agaccaggga	tgaaatcaag
19621	gaaaaaacga	acttgacgct	acctcttccg	ccttgagggt	gatgactctg	gaggtgggga
19681	tcttcctttt	gggcttcgcc	ctcagtaact	cctcatagtt	cttactctct	gccccatttc
19741	cataaatcat	ttgtagcttc	caggtggctt	cctccactgc	ctcatccctg	ttccacgcat
19801	ctgcctcttc	tctgctcagc	tcttcctgct	tatgcaggag	tttggtcagg	ttcttcagct
19861	ctccctccca	ctcctgccaa	ggcagagtag	ggccaagccc	ctcttgacac	agacacagag
19921	aacaccgaca	aacagagggg	cccattgccc	cctgtggagg	gccagggggc	tgcatattca
19981	gaaaatctag	gtcatctatc	tggggtctcg	aagaggggaa	tggtctgccc	aaagtctcat
20041	aaaaatgcaa	ggaccaggct	agacccctc	cccagctcac	ttgatgaact	ttctcctggg
20101	atgggctctt	ttcattcaac	aaaggcatga	gctgtatttg	caaggaatat	aatgttatct
20161	ctgagagaca	gtgagaagga	aagaaagaaa	aggaagaaa	aaaggaagag	aaggaagaa
20221	gaaagaagg	agagagggag	ggaagaggaa	ggaaggaaga	aagaaaggaa	ggagggaacg
20281	aaaaaggaag	aaaggaagga	aaaaggaagg	aaggggaagg	aggagggagg	aaggggaacac
20341	ttacctggct	agacagaagg	tggatttttg	cctcactactg	cacacagcag	ccagagctca
20401	cttgtacaa	gcaggaagta	catatgcttt	ctccagacac	tggtagaaac	attgcttgct
20461	ggatgatggc	attgatcagg	gagctcttcc	cagccccagt	gcttccaaat	aatgcaatgt
20521	agattgggtc	cactgtcggc	ttttcaatca	aggcaagaag	cctatttctg	gatgaatttt
20581	aaaatgcaca	ttgccatgat	catcaaagtt	tagagtgaag	agcattgaaa	cagaatgcaa
20641	tgaactgaa	ttctaattccc	gggcacgtca	ctacaacctg	agtctctgtt	ctctcctctg
20701	acaaaggaga	atatcaatcc	ctgttgacca	cacagagtgg	gcctggggat	caaatgagat
20761	tatgtaaaaa	tatttggtaa	tgataattcc	acacaatccc	tgtaatgtct	atataagatg
20821	gaggtgtctt	tatttgattt	cactggaaat	tgcatattct	tttttttggt	tgtttatgtg
20881	tctgagacag	gatctcactt	tgtcacccag	gctggagtgc	agtgggtgcaa	ttatggctca
20941	ctgcagcctt	aacctcctag	gctcaagcga	ctcccttgcc	tcagcctctg	aagtagctgg
21001	aactacaggc	gtgctccatg	acgcctggct	aatttttttg	tattttgtat	agagacagag
21061	ttttgccata	ttgtctcaggc	tgctctcaag	tggtccaccc	gcctcggcct	cccaagatgc
21121	tgagattata	ggtgtcagcc	actacacctg	gacagaaatt	gtatattctt	gacatagaa
21181	atacaatgta	tacagacatg	aagtattaat	agtgttagga	gcaacacaga	gtcttactga
21241	gcaggagaaa	ttgtagcctc	agcaccactc	gagtccttca	agtcagcaga	ggacaaataa
21301	ttggggacat	gagctcagag	tcacacagcc	tggttggaat	cttggaacct	ctcactgact
21361	actggtgtga	gtttgttcac	atctgggccc	ctctgagcct	caatttcagt	aaaatggaa
21421	tgatcataat	agctcctact	ttataggagt	gttggtggg	ttaaatgata	tttgacatat
21481	ataaagcact	tagcagatta	cctaacagtg	acagctcaat	gaatggcaac	tatcccatag
21541	ctacagaaat	ttgtgtccct	cttttcttag	tttctaatta	aatagtcctt	ttccttggtg
21601	ggggaagaga	cctataatcg	cctaaggaag	gacactgaca	taaaatttca	gggtgcaga
21661	tttcagactc	acatgagata	cttgactcca	ttagggatgc	tgtcatccag	gaagacagac
21721	tgaataagtt	tctgataagt	gttgctcaaa	acccttctgg	tccgtgattc	caatttttca
21781	tctggaatta	acagagagat	gtcacctgca	ggtacctggc	agggataaa	tactaggtct
21841	ccaaaggagg	tgtctgttga	acatcgcaga	gaagggaaga	aaaaaatgaa	atgaaaagcc
21901	gattaaacaa	gatagcaggt	tatctcacac	gagggatggc	ggctcactgg	agatagcaat
21961	cattggttct	gaaaatctcc	aaagagccag	gctgctctga	aactctggta	acagaagtta
22021	tggtatgctt	ggccacgggc	ttcagccctt	tctagatttt	tcctcctctg	ggtcataaaa
22081	agaaagaaaa	aatgcaagga	caccatgctg	acaccacagg	aagcaccac	ctcatcattt
22141	tcacaggctt	tgttctaate	tgggaagaag	tattttgata	acaacctctc	tgcccacaga
22201	tgcttgagga	caaattgcct	aactgcatac	cgctacagag	actttgagcc	ctctctttcc
22261	aaaacagggt	tatttgatgt	ttcatccagc	ttcagaattt	aacctcctct	tcctttgacc
22321	atttttaacat	acttggaaca	gaaccgaacc	atgttaaaac	aacagttcca	agacactcac
22381	attccttaag	agcactctgc	tccatggagg	gaaatgctcg	gaaccgctgg	tctcgatctg
22441	atttccttct	ttttctcggt	cgttctttat	ataaatcatc	ttcaactaga	gataaacaga
22501	ggatatataa	gattataggt	gttggggtat	ggtattcttg	cctcacctct	atcagcatag
22561	aaactcaaag	ggtctaacag	actggctggg	tgcatgggct	catgtctgta	atctgaagca
22621	tgagggatta	catgcctggg	aggccaagg	gggaggattg	cttgagggtca	ggagttcgag
22681	acatgcctgg	gcaacacagc	aagacccttg	tctcaaaaaa	aaattttttt	ttttggtaga
22741	gacaggggtc	ttactgtgtt	gcccagcagg	ccagagtgtc	gcgcatgat	cacgactcac

22801	tgcaacctca	acctcccagg	ctcatgtgat	cctcccacct	cagcctctag	agtacctggg
22861	actacaggca	cacaccatca	cacctggcaa	atTTTTgtag	agacagtgtt	tctgtatgtt
22921	gcccaggctg	gtctcgaact	taccaaaaat	atatatatat	atattttttg	attagccagg
22981	tgtgggtggg	tgcacctgta	gtcctagcta	ctcaggaggc	tgaagtggga	gggtcacttg
23041	agcccaggag	ttgtaggctg	cagtgggcta	tgattttgcc	actgcactcc	aacctgggtc
23101	atacaagacc	ctgtctcaaa	aaataaaaaa	ataaaaaaat	aggaactgct	tccagctctg
23161	tctgcccata	aactcatttg	ggaaaaaaca	ctgttggtat	attccagttt	ggttcctact
23221	tctcagagtc	acaagactgt	ctgcatttag	ttccaaaatg	tagagttcaa	tctatctccc
23281	cagataaata	gtttctatatt	gcaactcctc	agaagccaaa	gtggctggac	attcttcacc
23341	aaatgccaa	tcttaggaaa	tcaaccaact	gtgtagagaa	aagctgaaga	cgggggttaag
23401	tgtgcagtga	ggaaaggcct	ctgggaggtc	tttccagact	aggaaagaaa	gcgactagaa
23461	gcagaatgca	aaggaaaaaca	ggggccagggt	gaccagaaaa	aatcacatct	tcttgagcct
23521	caatTTtattc	atctggaaaa	tgagagaact	gacaaaaaaa	atacacatgc	attacaagag
23581	tttggtatgga	tgcccttcag	agatccttca	aaattttggg	aaattttttc	ccttcttcac
23641	ttggagacat	cagcagttgc	ttccttgcat	aataagcaag	aacagaaatc	acctagaagc
23701	agatctttgt	agctctgcgg	agatttcattc	acttaagcag	tttcccttgg	ttatgaagtg
23761	aagggtctgt	tgggcccatc	atcctaaaga	aggggtataa	ccaaaatgtt	gtgttttaga
23821	tggatttttag	aaaacacccc	ctggtcatgt	ttacctggct	agaaaggggc	aagaggtgga
23881	tctttgggggt	gaaggttggt	gaaaatgggt	ctaaccagg	gctcctgggg	ctggcctgca
23941	ccttagaaaa	tggccatgag	gtgaaccaa	agagtagatc	cccaagaatc	acagggaggc
24001	catgaggctc	caaaactgag	tgtagacttc	tgaggtcatt	ccctgcagaa	tcacagagaa
24061	gaaatgtcac	gtggcatcca	agaacggcag	cccacacaga	aggcccagtc	tctgcctgag
24121	ggctccagac	acagatctcc	cagacaattc	aggcagccct	ccctctctgt	tcaaccttgc
24181	atcttccagt	agaaatggaa	gccaacgcac	accccgtaga	atattctagt	gcaagccatg
24241	ctatttgcca	aggcctgaat	ctctgccacc	tcattgtcca	cgctcaagcc	acaactgaga
24301	gttgagagag	gttatacttc	gaccttcccc	tctcagacat	tctaactgat	gattgagctc
24361	aaattcaaga	attcgccctg	gctcggtttg	aggactgcct	tagaactaaa	taccaagtgg
24421	ttgctggcct	ggtggctgga	gctggagagg	aagaggcaaa	gaaagaagtc	tgctggtcag
24481	gcccctaggc	cagctaccgg	ctgaagaaca	ctttaatctg	tgctcgaaaa	agcatctcag
24541	aactccagga	tgtgtccaca	acctgtctgt	ggaaagagtg	gaaagggatg	tagactttgt
24601	aactcaggaa	gttttttggg	taaacaaagt	ttgcctaato	tcttagccat	tgacttaaaa
24661	caagggccaa	gtccacacat	ccagggccct	tatttcaaca	ttgtaaggac	tggcaacgaa
24721	tctgataaga	gttttggaagc	ttatttctca	actaaagcaa	atgctaggaa	gtccaatata
24781	gcagggttta	aaactttttc	ttcaatttct	aaaatgcaa	actatctgga	aaggagaaaag
24841	tattccatct	tggttttata	ccatacgagg	gtagtatttt	tattattagat	tgtaaccaaa
24901	ataaacttgt	ccgaatttag	atcacagattc	ttattattct	tgctccatgc	tcaagccctt
24961	ccttgccctt	tccagcattc	tccaccgaga	aaccagctgt	tgggggtggat	agcggcagtc
25021	caaaaagcSt	agtataactgc	tgctgttctt	tcttgctgcc	acaagggtgtc	agttttgaaa
25081	actgatcccc	caacagcctg	agctgagatg	Yggtttcatt	gttgactgtt	gaaatttttt
25141	tctttttttt	gagacgaggt	ctccactctc	gcccaggctg	gagtgcaatg	gctgggtcac
25201	ggctcactgc	agcctcaacc	tcctgtggctc	aagcaatcct	ccgacctcag	ctcccaagta
25261	gctggaacta	caggcactca	ccaccacgcc	tggctaattt	ttttttcttt	tttttttttt
25321	ttccagtgat	gaggtcttgc	tgtgttgctt	aggctggtct	tgaactcctg	ggctcaagag
25381	atcctcctgc	cttggcctcc	caaagtccctg	ggattacaga	catgattcag	accattgccca
25441	gaaactgccca	gaaatcttcc	aaagcaaacg	atcccattcc	tctaggctaa	gctcaactga
25501	ggtgacccca	aagtggcagc	tcttccata	tgccgtgctca	ggcttgcttc	agatgctagc
25561	atcaccccca	ctaaccacaga	aacactgctg	cactttgccc	ctaaggaaat	ccatcccttc
25621	ggagttagga	aacattccct	cttcagttct	caaaacacaa	aggatatgat	acgccactgt
25681	ttctgtcaac	tgggggttat	atccctgtca	tcgtcctata	gggaagcctc	ccttattggc
25741	tocaggggcg	caataccagc	cttctgaggg	agggaagcca	cccgcactga	ctccaatgtg
25801	agtcaccctt	ccaggagccg	ctccaaggca	cctacatctc	tcttctacg	gaaagccgcc
25861	taacgtcatt	ccaggggtct	tgtaagacag	gctacctgac	ttggacagag	agtccaagct
25921	gactttctgg	aacattatta	ttattgagat	ggagtctcac	tctgtcgccc	aggctggaga
25981	gcagtggcac	aatcttggct	cactgtgacc	tccgcctcct	gggttcaagt	gattctcctg
26041	cctcagcctc	ccaagtatct	gggactacag	gcacacacca	ccacaccggg	ctaatttttg
26101	tatttttagt	agagatggga	tttcacatg	ttggtcaggc	tggtcacgaa	ctcctgacgt
26161	catgtgatcc	accgcctca	gcctcccaaa	ttgctgggat	tacaggtgtg	agccaccatg
26221	cccagcctgg	aacttttttt	ttttttttta	aggaaaaagc	cgcacttatt	taactgctgg
26281	ccaaacagtt	cacttcatta	tgaacaggtt	atttcagtta	ctgttgaaat	atctgatctt
26341	cattcactat	ttaaacttgt	cacgagaaag	caccttctga	ataaagagta	aggaaacatg
26401	acaacagtga	aatatagtc	caacctattt	gccttttata	ctatatcagt	aatcagttcc
26461	tatatcctcc	atcagaagat	actgcaaaga	tgttacctgg	atgctggttc	tggccaaaaa
26521	catccttcgt	ttctgccatt	ccttggttacg	tgaactcctg	ctcttctcag	ttcaggagaa
26581	ccagcctgtg	aagacaaaagt	acaaagacag	gtggggcagc	ggaatttgaa	aatacagaaa

26641	gcaaaccccta	aggatgtgtc	agtctgcttg	tgcggtattt	ctgcccattc	caacgtatct
26701	ctgaatccct	cacaacacct	agtacactac	ttgatggaaa	tataacaaag	attagtgttt
26761	gaacactggg	ttccaataat	acttaggcc	ggcgagtg	ctcatgcctg	taatcccagc
26821	tctttgggag	gccaaaggcg	gaggatcact	tgaggtcagg	agttcaagac	cagcctgacc
26881	aacatgggtga	aatgccgtct	ctacaaaaaa	tacaaaaatt	agccaggcgt	ggtggtgggt
26941	gccataatc	ccagctactg	gggaggtga	ggcaggagaa	tctcttgaa	ctgggagcg
27001	gggctttag	tgagccgaga	tggaaacct	gcactccagc	ctggatgaca	aagcgagact
27061	ctgtcttgag	aaaaaaaaag	agtctaaacc	agtattacct	gagagcttta	gggtaactca
27121	cgctgaaatg	agaataagcc	tctggtgacc	agagctgtga	aggtgaattt	atatagcttc
27181	gctgcttttg	gtggcatcag	tgctccactg	gggcaaggct	tgaacctga	gccagtgtca
27241	tttagttcca	gaaagtctca	gcaggaatct	gggaggtgga	acattagata	ttcctagagt
27301	tccctggagt	tgctttgggg	ggctgctgac	tggggaagga	gatgcagtga	cttggttccc
27361	tggaccttcc	tcaccaagcg	caacttccac	caggaggaca	cccctcctgg	gctctccttg
27421	cagcttgtag	catggagcat	ctgcaggagg	gagaggagag	gcacatccaa	aagaagaagg
27481	ggcagatcca	gttgggtctcc	agtccaccag	cagatagggc	ttggacatag	agggagaact
27541	ggacccgtgg	aggtcagaa	cagagggcac	tgcccagaca	gggtagaatc	tgctagcttg
27601	gagctgcttc	atctccaccc	atagagaaga	atlttgaggaa	cacgagtggg	gaacacccaa
27661	ggtgcttact	tgccgttacc	tgatgtttct	gtttggatca	ctcagccgtg	tatttcaaaa
27721	gcagagtcac	cagcccttct	gaaatatgca	aaatcacctt	tgatctgtcc	tactagacct
27781	acaggctgtt	atlttagggaa	actatgtctc	atltacacat	taaccccaa	ttaaaatcac
27841	cattcatcag	ttactcttag	tttgggttgc	agatatatac	tgactaaacg	gcttctgtat
27901	ttggtgggtg	ttgggtgaat	ggggtgagaa	cccctattaa	gtgcaagatt	ctgagtctta
27961	ggtagagctc	tttccctctc	ccttccctcc	ctctgcagcc	cattttcaga	tcacccaca
28021	ctccttacag	ttctccaaag	caggggtatcc	tattgaactg	ctgagcattt	tagcacatgt
28081	ggggtgccct	tccttcttcc	tcgcagggtga	attcctgtct	ttgcttcaaa	atcctgttct
28141	agtgtttcct	ctgtgaaacg	gtgtctgatt	ccactcatt	ccccccgtgt	gtccccggt
28201	ccatacacac	ctgtgttata	acactgactg	tgctgtgctt	acgtctgttc	ctcctggttg
28261	aaccactcga	ggtcagagga	ctgcattttt	ttcagtgggtg	catttctatc	aagaagggca
28321	atgactaact	aaacgttagc	tgaatgaatg	gataacataa	tggatgggaa	gcttttgatg
28381	agattccaca	ccaaaggatg	tttaaaagct	aaatttggtg	ttttggtctg	gtgggttgtt
28441	tgctgttccg	agagtcctgc	tctgttgccc	aggctggaat	gcagcggcac	atcctcagtt
28501	tactacaacc	tcacctcctc	gggttcaagc	gattttcctg	cctcagcctc	cagagtagct
28561	gggattacag	gcattgcacc	atcatgcctg	gctaattaca	atlttagtag	agatggggtt
28621	tcaccatggt	ggccaggctg	gtcttgaact	cctggactca	agtgatccac	ccacttcagc
28681	ttcccaaagg	gctgggggtta	cagggtgtgag	ccaccatgcc	tggccctgaa	ttcatttttg
28741	gtatgttaga	gtccttgatt	gaagagtaga	gcaaagagta	gaaataaata	gtcaagtatg
28801	caaatgaagc	Rgcactcaca	gtgaggactg	ctccacgggg	atgagggaata	tttgatgtga
28861	ggWagggaag	acacaggtaa	aggcagaaga	tgggggcagg	ggccttcattg	gtggtcttca
28921	aatgcattac	cttacagctg	caaataaaag	agacatcctg	catattttct	agggttccca
28981	gagggtagaa	atagatccac	taggtggcag	gtacaagaaa	gcattttggt	ttaacattct
29041	caactaatca	gagttgtcca	aaatggaaatg	agctccctta	ggaaaccgtg	agtttcttgc
29101	ctctagaggt	gtctgaacct	atattgatat	gaccaccta	taacgacatt	gagatagata
29161	aattttaagc	cttcttcaaa	ttccaagatt	ctatgttctt	acttgaaatt	ttaccaaga
29221	aagctggtag	tagaagtgcc	ctatgaggtc	accaaatttc	acaagaggca	caagcaattt
29281	ctggtttgga	atactgagtc	agaaatgtct	acacacatca	aaacctaggt	gagctacaag
29341	aaaaaaaaag	gtagatgggc	tataacatga	gtaaatatac	cataatgcat	gtcaataatt
29401	ttaggcaggg	tgtagtggtc	catatctata	atcccagcac	tttggggaggc	tgaggaggga
29461	ggattcattg	agccctggag	tttgagacca	gcctgggcaa	tcaagtga	ccttgacact
29521	acaaaaaa	aaaaaaaaaga	aagaaagtag	ccggttgtgg	tgacacacac	ctgtagtacc
29581	cactatttgg	gaggctaaga	tgggagatc	ccttgaaccc	tggagttcaa	ggttgagctg
29641	agctatgatt	gcaccactgc	actccagcct	gggtgacagc	aagaccctgt	ctcaaaagaa
29701	aaaaaaaaatt	agatacagaa	ctaggattac	ccctaggaaa	aagatctgtt	gttttctgaa
29761	gagagcagct	taaggagtaa	ctcgtatgaa	atgctaggca	tcataattatc	acagttagca
29821	ctggaagcaa	aacaaaatgt	atgatacaac	cagtgaacaa	atccatgggtc	gacctccagg
29881	acagctgggt	tgctggaatt	caaattctag	gaggaaacag	tggagccgga	taaagtacaa
29941	aagaactgat	agaccaagtt	gataagagtc	gtggtcatac	agacactaag	attaggagacc
30001	ttcaatttag	agagctgtgg	attaaagggc	acaaaatcaa	agggccttga	gttttagagca
30061	gaagcagctg	tggtcactga	attctggaat	acgggtagta	gttttggtct	ctttcagaat
30121	cacaagaact	ttagttcatt	tgacttcatt	tactaaatca	acagtacaca	gtgtccagca
30181	agggccgtta	gcatgtaact	agctgggggt	ggtttgggtt	tagctgcatg	caatgaagac
30241	agaaaaccac	acaaattcct	ccgtagcagc	atctatgggt	tctgtatata	agtaatttgt
30301	gaactctaac	tgaaaatttg	aatcattgga	aattatgtgg	tttgggtgtg	gtggattggg
30361	cttcaaattc	ctttggccaa	tgttcaaaga	ctggcaggaa	gggatcattt	gatgacaaa
30421	caaaaacact	agcaaatgga	aagggatcat	ttgcaaaaag	gctaagagga	cagaagaaaa

30481	taaacaccaa	gtacataaaa	aggaaaagcg	ataaacaat	gggggtagtt	tggtagagac
30541	attcagtata	tagtcattgg	gttgaactaa	atgatgaggt	gggaagaaca	ttaggctgga
30601	aatctggtgt	ggaagcttcc	agctccagct	ttaccaccca	tcaataaagc	tggaggactt
30661	aagacaagtc	gttttatcac	cctggcctca	ggtttcccat	gtgtaaggca	aagactgaac
30721	tagtccttaa	agtgtgagtg	ttctctgatt	tagatgggtg	ctaaaaagac	acaaattctc
30781	attttgaaaa	tgtcacctgt	cattgaaaMt	gagaccgtac	caatgattag	ggcttccgtg
30841	tggccaagag	cagtagacca	agagcactgt	gatttcaagt	gcattgtatt	aaggagaaaa
30901	tgagctacag	gagtatagtc	cttcttcaaa	tcatacctgca	gtctatttcc	cccccagggg
30961	aaggaatggt	ggtgttcatt	gtgaaaaatt	ggtttctgtg	gagaagcaat	tgataacagt
31021	ccaaatttat	tgtcagata	agcaagggtc	tgctgtgtgt	gtttattttg	cttgattttc
31081	agctatgtgg	ttagcactga	agttttgctt	ttagtgtcta	tttacattca	tgtgactga
31141	acctaagaat	agcatctttt	aaataagact	gtgcacattc	agctagatat	aggaaggctg
31201	cacagaactg	atttctgat	aaataattgt	aacttttcac	agttattctt	gcttctctcc
31261	ccctcctttt	cttctgtttg	tgtttgctta	ttgttttttt	tcttacagga	aatcacaga
31321	ggcaacacaa	acgcagggtc	ttcctgtact	agaacagctg	catatcatgg	cattgaataa
31381	agaatgcggg	ccgggtgcgg	tggtctctgc	ctgtaatccc	agcacttttg	gaggctgaag
31441	cagggtggatc	acaaggctcag	gagtttgaga	ccagcctgac	cagcatgggtg	aaaccccatc
31501	tctactaaaa	atacaaaaat	tagcccgggtg	tggtggcacg	cgctgtaat	cccagctact
31561	caggaggctg	atgcaggaga	attgcttgaa	cccaggaaat	ggagggttgc	gtgagccgag
31621	atcacgccac	tgccagcctg	ggcaacaaag	cgagatttgc	tcttaaaaaa	aagaatgcta
31681	gggcccttcc	ttctacagat	ccagaccctg	aggaccagga	aggggaagaa	acttgtccaa
31741	ggtcacatag	taagattatg	gcaatgaggg	ctgaaactca	gttttctctga	ctccaagttc
31801	aatgttcttt	tcatacattt	cagcagcctc	taataggcac	tatttaaaca	agccccctgg
31861	agctcaggca	tagggcaatg	gaaggagtag	tgcccgc aaa	ggagggttcc	ctcaaatgca
31921	tggaagggc	atttcatgag	aaggaaggct	gttgagtggg	gagagagagt	agcaacttga
31981	tagaattgta	tgcaaggaat	gaagctgaa	ctactgcctg	caaactgctt	ttttgttttt
32041	gtttttgttg	ttctttgttt	ttttgtttgt	ttgtttgttt	gttttgagat	ggagtttttg
32101	ctctatcgcc	caggctggag	tgcaatggca	catgatctca	gctcactgca	acctccgcct
32161	cccaggttca	agcgattctc	ctgcctcagc	ctcccagta	gctgggatta	caggcacctg
32221	ccactatgcc	cggtctaat	ttgatctctc	agtagagaca	agggttccat	atgttggcca
32281	gctgtgtctc	gaacccctga	cctcaggtga	tctgcctacc	ttggcctccc	aaagtctctg
32341	gattacaggc	atgagccacc	atgcccaacc	ttactgcctg	caaacttgaa	gacacccctc
32401	gaagtaaaac	aagccagaga	gagagagaga	gactatcaga	ggaatacatt	ttcactgcat
32461	tttttctgtg	ttcaatgcaa	aattcaatgt	gtttattaca	aatgacaaat	gtccttttct
32521	tttcccctag	ttattctcct	taggcagaat	ctcatttcac	cttactctgc	ctcgcatcat
32581	taaccccttg	caacactatt	tactctctta	aattttaaag	ctttaaattg	caaagtttct
32641	caaaggggca	tttttgtcat	taacattggt	ttactctcta	cttttaataa	atatgggaag
32701	tacttggtac	acagcaagca	ctaaataatt	tgttgaatga	atgatgatag	atataatttt
32761	ttgtatctat	atctagatag	atctatctag	atctagatat	agatatatat	cttagatggt
32821	aagatcaaga	atagcagaaa	ttatatctac	tattttattt	atccatccat	tcattcaaca
32881	aatattttgtg	agtgcctatg	atgcaccagg	cactgttcta	gatactgggg	gtacagcagg
32941	aaacaatgtg	gtcacagtat	ctgcaaaaca	agggcttcca	tattccagat	gtgaatgggtg
33001	gagcaaaaa	tattagggat	ttggagtcag	aatgcctgga	tttttgctcc	tgctcccaat
33061	cttcaatgca	acacagagca	aatcacctga	tctctctgag	cctctatttc	tccacatata
33121	ggatcaaat	ggttttaaaa	aaaataacctg	tatcaagaat	tcatttggcca	ggcatgggtg
33181	ctcatgcctg	taattctagc	accttgggag	gccaagggtg	gccgatcact	tgaggtcagg
33241	agttcgagac	cagcctgggtc	aacatggtga	aacctgtctc	ctactaaaaa	tacaaaaatt
33301	acctgggcat	ggtgcagcat	gcctgtaatt	ccagctgctt	gggagactga	ggcaggagag
33361	tcacttgaag	ccaggaaaca	gaggttgcag	tgagttgaga	tcgcaccact	gcactccagc
33421	ctgggcaaca	gagcgagact	ccacgtcaaa	aaaataaaaa	ataaaaaaga	gttcgtttgtg
33481	agaattaaat	gtgataatga	atgcaaaatt	gctctgcagg	cacttatgca	tctgtacaaa
33541	tgttattttt	tgaatatttt	atgaaggtag	ttaaactaga	aatcagtact	agtgtcatcc
33601	aggccacgtc	tgtagtctt	tggggaacag	atgccacctt	agctgagatt	ctgactcccc
33661	ttcttgtccc	tattgatcaa	acacattgag	gatgctctac	catcttctct	aaagaatctt
33721	cgatgcagtg	ggtgcagtg	atcacacctg	taattccaac	attttgggag	gctgagccgg
33781	gcagatcatg	aggtcaggag	tttgagacca	tcctggctaa	ctcgggtgaa	ctctgtctct
33841	actaaaaata	caaaaaatta	gccagccgtt	gtggcacatg	cctgtagtcc	cagctactca
33901	ggaggctaag	gcagaggaat	cgcttgaacc	cgggagcgcg	aggttgagat	gagctgagat
33961	catgccactg	cactccagcc	tgggcaacag	agtgagactc	catctcaaaa	aaaaaaaaaa
34021	caatcttttg	tcaagttttt	tcttgcctc	acaagaagg	gagctgaaga	atgaaagaag
34081	gaagaggagg	agactgttaa	tgaatgata	gaatgttgat	gaaatttaga	atttatttcc
34141	tttctctttg	agtttgggtt	taggtttaag	taactttggt	gttggcaaca	tttcttaacc
34201	acctcggttg	ccactgcact	gctcttccaa	actagggatg	tcccaaggca	gcaatttaaa
34261	cacattagct	cagaagcaaa	agtgaacaa	gggcctccaa	atctctaaag	gaaaagaggt

34321	atgtttttcag	cattaacaca	tcggctaagt	tgttcagaat	tagaaatgtg	gccaccagaa
34381	gcagtcgtag	tagaaggag	gcttttaggat	gacatgtttt	tctccatctt	tcttccctag
34441	gtaaagaggt	aagcaaaata	agcagatatt	ttcataacgt	acttttagcag	attctagcga
34501	aacgaatttg	aatctttctca	tattgtctctc	ccaagcttgt	taaagctaata	gaggcataag
34561	atgacacttc	cctcctttatc	aggaattcat	acacctgcc	gcaatttctg	caatagctgt
34621	tttcttttgg	gaggaatgat	taatgcattg	cctaagatca	cactgttctg	taaaattcaa
34681	actaaaacct	gatttctccc	atttctagcc	cactggcctt	tttaccatagc	cttaaaattc
34741	acttaaacag	cctgagcatg	gtggctcatg	cctgtaatcc	cagcactttg	ggaggctgag
34801	aggagagggg	atcccttggg	cccaggagct	tgagaccagc	cttggcaaca	aagttagacc
34861	ctcatctcta	caattttttt	tttaattagc	tgggcatggc	agcatgtatc	tgtagtccca
34921	gttattcag	aggctgaggc	aggaggatta	cttgagccca	gaagggtgaag	attacagtga
34981	gccatgttca	taccactgca	ctccagcctg	ggtaacagac	cgaagcctca	aaaacaaaca
35041	aaccaaaccc	ccaaaattca	cttaagcaaa	cagaaaagta	aaattcactt	tccctcaaga
35101	aataacttgc	tttaagaaaa	tcaaaggaga	gagaagagac	aaatgtccca	tatacagaag
35161	aattccaaat	aattttacgta	gctactccat	cctcaaggag	gtagaataaa	actccttgc
35221	ccttaagtgt	gtgctgtgca	tagcagcttc	tttccaaaga	gaacacggcg	gagattggag
35281	cagggaagg	atcaccttat	agcagagaaa	ccagacaagc	gctgcctcta	ccagggtggtc
35341	aagggtcaaca	ccaacagtcc	taaatcatgt	tgacaatatt	atcccttgat	atgatgtgat
35401	gagaacagca	ctttaactct	gggactcctt	ccccaaaacc	cataactcca	gtctaataat
35461	tgaaaaaca	tcagacaaat	gccaatagga	gaagcctaca	aaatacccaa	ccagttactcc
35521	tcaaaactgc	catattcatc	aaaaacaagg	aaagtctgag	aaacaacgct	acagttaaga
35581	ggagcctaag	ggaacataac	taaatataat	gtgctatcct	gataggatgc	cacagtagaa
35641	agcagacgtt	agataaaaac	caagaatctg	aatcaagcac	agactttagt	ttaagaataa
35701	cgtatcagta	tcggttcatc	agttgtaaca	aatgtaccac	actaatacaa	gataatggg
35761	aagttgggtg	ggaattacat	gggaatgctc	taaactatct	tctcaatatt	cctgtaaact
35821	caaaactgtt	ctaaaaatta	aggtccattt	ttaaaaaatc	aaccacaagt	tacagtaaca
35881	attacataaa	agaaacaact	taccagatgt	cactgaacag	aaaaaaaaaa	aaaaagtctt
35941	ggtttggtta	caggagtcca	cagaggagat	ggtggcagct	tctccttgg	tcagggtattt
36001	ggaaccagct	ggggtacgta	gccttttggg	ccacaatcac	aaagccacgt	ttacacagca
36061	gaagtgaac	caacttgcca	tcagggagct	atttaccatg	gtctcctccc	acagatactg
36121	atattattctg	tctcttttca	cttaaatcta	aattcaccag	gcacgggtggc	gtactcctgt
36181	aatcccagca	ctttgggagg	ctgaggcagg	cagatcactt	gagggtcagga	gtacaacagc
36241	agcctggcca	acatggtgaa	accttgactc	tactaaaaat	acaaaaaaaa	aaaaaaaaatc
36301	agccaggcgt	ggtggcgaaa	ccctgtagtc	ccagctactc	aggaggctga	agcaagagaa
36361	tcacttgagc	ctgggaggtg	gaggttgcag	tgagccaaga	tcgcgccact	gcactccaac
36421	ctggatgatg	gagcaagact	acatctcgaa	aaaaaaaaaa	atctaaatgg	tacatgaaac
36481	ggggatcttg	cacctttctc	tgccatgttc	ctacaagtag	acagatgtct	tttttctctt
36541	ccaagtgc	cactgcatcc	ttccacagca	tcttttctct	gagggtgaac	aatattcaga
36601	tgactgggtc	actggagcag	acaggaaaaa	tgccccagtt	ccaagacaga	caaagcactg
36661	gtccccgac	acccaactat	aagacatgac	tgaggagatg	ggaggatatg	agataataaa
36721	tataaagtta	ttcttgaccc	aaagtcttaa	aattcaccaa	aaataacatt	tgccaaaaat
36781	aaaacaacat	tggattgttt	gagtgtagaa	atgagtgata	taccgcttta	cacctctttc
36841	tggaaacttaa	attttaaatc	cgtatatttt	tatgtcttta	tttagttaga	gacaaggctt
36901	tgctctgtca	cccagcctgg	agtgcagttg	cgtgatcata	gctcacagca	gcctcgacct
36961	cccaggctca	agccatcctc	ccacctcagc	ctcctgggta	gctgggacca	tggctcatgct
37021	ccatcacgcc	catctaattt	taaaattcata	tttttaataa	accagaaaaa	caaaaatttg
37081	aaagtgtgtg	ccaggagaga	agaaaaataa	acccactcca	ctttgctgtg	tctggcaacc
37141	caaaccatca	catcacatac	agggaaaaag	ttgtcccttg	aacctcagc	ttcagggtgt
37201	gtgttacctt	tttttttttt	ttgtgagaca	cagtcttgc	ctgtcaccca	ggctagagt
37261	cagtggcgcc	gtctcagctc	actgcaacct	ccgtcttcog	ggttcaagag	attttcctgc
37321	ctcagcctcc	tgagtagctg	ggactacagg	ggtgcgtcac	catgcctgac	tttttttttt
37381	tttaatttta	gtagagacag	ggtttcacca	tgtttgccag	gctggtctcg	aactcctgat
37441	ctcatatgat	ccgcccacct	tggcctccca	aagtgcctgg	attacaggta	tgagctactg
37501	tgctgggcca	agggtgtatg	ttacctttaa	aggccaactc	agtcattacc	aagacaaccc
37561	tcagtactct	aggtatgtat	ttagtctgaa	gaggggatcc	agaaggtaaa	gaggtcagaa
37621	cagggaagaa	attgaagata	aatgtttaga	ctgtggtctt	aaaaactgct	ttgaactttt
37681	ggcaagatga	cgtccctcct	ccataagaaa	accacagaaa	caccaggcta	aagttcagaa
37741	caaagaaagt	cagcacaagt	gccatcctgg	atgctgagtc	ttatactgat	ctctctaggc
37801	aaaatcacac	cttacgctta	cattcatgtc	acgtgtgctc	agctctccct	ctactggcta
37861	tttcttagct	tcaacagtta	accacaatag	ctactaaatt	acttaattat	taaaattaat
37921	tttgagcatt	tttaaagctt	cctggggcaa	ggcccagtg	ctcacacctg	tgtacctg
37981	attttgggag	gccgaggtgg	acagatcacc	tgaagtcagg	agcttgagac	cagcctagcc
38041	aacatggtga	aacctgtctc	ctactaaaaa	tacaaaatct	agccggtcat	ggcagtggtg
38101	gcctgtagtc	ccagctactc	aggaggctga	ggcaggagaa	ttgcttgaac	atgggaggca

38161	gaggttgag	tgagccaaga	tcaggccact	gcactccagc	ctgggcgaca	aaggagact
38221	ccatctcaaa	aaataaaaaa	ataaagcttc	ctgagttcct	ttatttggtt	gcttggtttt
38281	gtttgtttgt	tttagacagg	gtctcactct	gtcaccagc	ctgaagtgt	gtggcgttat
38341	cacggctcac	tgcaacctca	acctcctggg	ctcaagtgat	cctcccacct	cagcctccca
38401	agcagctggg	actacaggca	tgtgccacca	cacttagctt	aaaaaaacaa	aaaacaaaaa
38461	aagaaaaaca	aaaacaaaac	aaaactttag	tagagatgag	gtctccttat	gttaccagag
38521	ctgaactcaa	actcctgagc	tcaagcgatc	ctcccatctt	ggcctcccaa	agtgttgagg
38581	ttacaggtgt	aagccactgg	gtccagcctc	tgggttctt	tatggtagtc	cttggcacct
38641	gctattggac	gcctgaatgc	taggcctggt	agatgtttct	ccattctctc	aggctgaatt
38701	tccattgagc	tactgaaaa	cactgaagca	atgagaccgt	ccacagagcc	tcatcatcac
38761	cccccaacc	ctatccgggtg	ctctgcctc	cagcctgttg	ctagatgaac	tatccatgca
38821	cccattaaag	gccaatccct	tctctgtgca	ccagattcca	tccctctctc	taactgaaga
38881	accatgtaga	gtacttctct	ctcttctcta	attcaccaaa	ctttcactct	ccactggatt
38941	attcccatct	gtgaacagat	ttgatgtcat	tcatccacc	ttaacaaaac	ctccctggac
39001	ctccatttcc	ctgccaggta	ttagtatat	tctctgcact	tctttatgac	aaaactccta
39061	ggctgggcat	ggtgtcagca	ctttgggagg	ctgaggcagg	aggatcactt	gaaccaggga
39121	gttcaagacc	agtctgggca	acataaggaa	acccctgttt	ctacaaaaaa	taaaaataaa
39181	aattagccag	gcattggtgt	gcattgcctat	agtcccagct	acttgagggg	ctgagatgga
39241	aggatcactt	gagcctggga	ggtcagggtc	gcagttagct	gtgattgtac	cactgcactc
39301	cagcctgggt	gacagagcaa	gaccctgtct	ccaaaaaac	aaacaaaacc	ccacaactcc
39361	tcaaataggat	tgtctctact	tactgtctca	aattccttct	ctctcaagct	aatataaac
39421	tattccagtc	aagccttcac	ccttcccatc	ccattaaagc	tgttcttgtc	aaagtccaca
39481	atgatcctgg	tcaattttca	acctttatct	ttcttgagcc	atcaggagca	tttgacctgg
39541	ttgatcattc	cctcctgttt	gacaaacctc	ctacacttgg	ctttcagata	accactcccc
39601	tagttttcat	cccatctccc	tggagtgct	cctcagctct	cttcactggt	tattccccct
39661	cttctcaacc	tgttaataat	aaaatacccc	aaggcttcat	cttcggtctc	tttcttttcc
39721	acagccaccc	tgtttcgtag	ctttcaattt	cgttcacata	ccgatggctc	tgcactgata
39781	tccagtgtca	acctctttcg	tgtcttctcg	ttcactagtc	atccgaaaat	acaagttcaa
39841	accaatcccc	gccagtcctt	tgaacacctt	ctccacctaa	ttttctccat	ttcatcta
39901	gataactaca	tttttccagt	cccttgggtc	aaaagctttg	gtgtcacatt	tgtgtgtgt
39961	gtctgccttt	catattccac	atctgatctg	tcaaaaagtc	ttgttgaaat	cttcaattta
40021	tattcagaat	ctgaacactt	ctcaccacct	tactgtctga	ctaccccgat	ttgagtctca
40081	ataatctctg	gcctcattca	gtggttctca	agttttgctg	cacgttggaa	taaccaggga
40141	tcttttaaac	atgctaattg	ctgactccca	ccccttgata	ttctgattta	ataggtgtgg
40201	gatgtaactt	gggcactggg	aatttttcac	tgctctccag	gtgattccaa	ttgcagcaaa
40261	gtttgggaat	catgggctg	gctattgtga	ccgccacca	cctgatctcc	ccactccac
40321	accaaccccc	tcccacagtc	tattctcaat	gcagcaaata	gacatgcttt	taaattacag
40381	atcagatcca	ttcaattctc	tgctaaaaac	accagtggct	ccccatctca	attagggtta
40441	aagccaaagg	cttttcaatg	gcccaacaag	tgttacatga	gctgcactgc	cccctgtccc
40501	atgcctcact	cctctgacct	ctcttccaca	ttgcctcacc	cactgctccc	gtgatgtcag
40561	cctcctcagt	ctccttgaac	acccagaca	tcctctcacc	taggactttt	tttctttttt
40621	gagatggagt	ctcgctctgt	caccaggcta	gagtgcagtg	gtgtgagatc	taggctcact
40681	gcaactgcca	cctcccaatt	ctcctgcctc	agcctccgc	gtagctggga	gtgcaggcgc
40741	gtgccacat	gccagctaa	tttttgattt	ttcagtagag	acagggtttc	accatgttgg
40801	ccaggatggt	cttgatctct	tgacctgtg	atctgcctgc	ctcagcctcc	caaagtgtct
40861	ggattacagg	cgtgagccac	tgcgctgag	cttcacctag	gacctttgca	ctaaccatta
40921	tcttagccag	aaatactctt	tcccacatat	ctacacatca	agtattagct	caaataatcac
40981	attgKcaata	aggcctactg	taacccttat	ttaaaattat	acctttaacc	acctatctcc
41041	ctcccctctc	tgagataaaa	cagggaagag	gtagaggatg	gtgccattag	ataaggaaga
41101	gcagatctgg	agtaataata	gagttctggt	tgggacatgt	taagtttgag	atgcctatta
41161	gacacccaaa	tctaacaaga	tatcaaatag	gggaaaagtt	gggttgagac	tggaaacctg
41221	ggagcaaaca	cgtggtgctc	tggggatcac	cttgagagga	caaggctgag	tgaccccat
41281	ggaacactag	tattctgagg	tgaggcagag	acagagtcca	caaaagaaag	agaagagaaa
41341	caaagagagt	tggagggccc	tgtcataaaa	gatgttcagt	gaagagtga	caaaaagagt
41401	caaactctgt	aaaatacttg	aagagattta	ttctgagcca	tgattgacca	actgagccaa
41461	gacaggcaag	ctccaaaatt	ggggctttgc	ctgggagggt	tcttggtttt	gcccaggaaa
41521	caattcaagg	gtgagctgat	ggtgttaaat	agcaacttgc	attgaagcag	cagtgcacag
41581	ctgcagcaga	gggactgctc	cttgcgagc	agggctactc	acaggcagtg	cccagaagag
41641	cagctcagag	gcagttctgc	agtcataatt	ataccactt	ttaactgtat	tcaaatttaag
41701	gggcaattta	gcgagaaatg	tcaagaatga	ggatggtaac	ttccaggtca	tcaggtcatt
41761	gcatgggaaa	ggggagggtta	atgttcagggt	gttgccacgg	caatagttaa	ctgacatggc
41821	atactggtgg	gcactttatg	gaaagttgct	tccacccctg	ccctgtttca	gctagtcctc
41881	aacttgatcc	agtgtccaaa	ctctgcctcc	agaacagagt	cccacttctc	acctcacatg
41941	gtctgtgaca	gagccccagg	ggatcctgag	aacatgtgtc	caaggtgggtc	aggttacagc

42001	ttgatttttat	acatttttagg	gagttataag	acattactac	atgtaagatg	tatatgtggt
42061	tgggtccagca	tggaggaaaa	gttaaataatt	aaatctgaat	tcaattgaac	ctggacacaa
42121	acaatagtc	ccaagtcctg	gaacaagttt	tgtgagtc	ttgaggcttt	catccagcgc
42181	tgttttcagag	aaatctctat	ttcaatctat	tcctatacat	tagttattga	aaaacaatag
42241	acaatagcaa	aaacaagttg	acctttttgt	gttccttgag	cctggttgtg	aaggggccctt
42301	gtgactgggc	ctcatgccaa	acaacttggt	acaaaaagag	ctagggtccc	aggcccagcc
42361	gaagcttcag	gagacctatc	ctcatctgtg	caaggaggag	tggccaactc	tggagcccag
42421	gctgttgctt	cctgggtctg	tgtgtaatcc	tccatagtct	ggtgagtgtg	gtgcccact
42481	ctggagccca	ggatgttgct	tcccggtctg	gtggtgaatc	ctccatagtc	tgggtgagtgt
42541	agtgtccaag	tctggaaccc	aggctgttgc	ttcttggtct	tgtgataaat	cctccatagt
42601	ctgatgggtg	ggtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgtatata	catatacttt
42661	cccttctacc	cttcccattg	caatttgctt	attatatctg	cattgccatt	tacatgggat
42721	aaaggtcggt	tacccctaaa	ggtattgtgt	atgtgtcttt	tcttctcccc	tcacgcatat
42781	cccacagaga	acaaccagaa	aggtgggaca	actcaaagca	ttgatgtggg	gtcagggggc
42841	aacactttca	ggtcacaggc	agattcagtt	ttctctgact	ggcaattggt	tgaagaggtt
42901	attatctata	gacctgaaat	gaacagaaag	gaatgtctgg	gttaagacac	agagtgtgtg
42961	agactaaggt	tttatcatgc	agaggaagcc	tccaggtagc	aggcttcaga	gagaatagat
43021	tgtaaatggt	tccttttttt	tttttttttt	tttgagccgg	agtctcactg	tgtcgcccag
43081	gctggagtg	agtgggtgcaa	tctcggtcca	ctgcaagctc	tgcctcccg	gttcacgcca
43141	ttctcctgcc	tcagcctccc	aagtagctgg	gactacaggg	gcccaccacc	acgcccggct
43201	aattttttgt	attttttagtg	gagacgggtg	ttcatcatgt	tagccaggat	ggtcttgatc
43261	tcctgacctt	gtgatccgcc	cacctcggtc	tcccaaagtg	ctgggattac	aggcatgagc
43321	caccgcgccc	ggccgtaaat	gtttcttatc	agactttaag	agtccgttct	gtgctctatc
43381	agccttaagg	tctctgtgtt	gatgttaaca	ctggttagca	gctcctgaat	tctaaaaagg
43441	aggagggaat	aaggaggcat	gtccaacccc	actcccac	atggcctgag	ctagtttttc
43501	agtttaactt	tggaaatgcc	tggctgaga	attattttog	gtttacagag	gagaacttat
43561	gttcattggt	ataggatgac	aatgaggcca	gacaaatacc	actgaacttg	gcaggctggg
43621	aagttacaaa	gcttagcaag	gccagtttca	gtgtcatagt	ggggtcaaag	cctcggtgtt
43681	gcctcaattc	tagtagggta	catattcctg	tctttattaa	tgggagaggt	tcttaatctc
43741	cagcccatgg	acaaaggaga	tcatagatgg	gtttcaggaa	aacatcccaa	gtcctgcctc
43801	caaattttgt	aaaattttgt	gcctgtgtat	ttttctggag	aacgtaaacc	aacactgttc
43861	tgagacaggt	ctcagtcctg	aggtttattc	tgccaagggt	gaggacacac	ccaggaaaaa
43921	gagacataag	ttatcatgga	atctgtggcc	tgtggttttt	ccaaagaggg	ttttgaggac
43981	ttcaatatatt	aaagggaagg	agcagacaga	aggggaagaa	ggaacaacta	tgcattcatt
44041	tcacactcag	taaatctgcc	ttttacagaa	gacaaagtaa	acatagagga	aggagtcaaa
44101	tatgcatttt	tcttgggtgt	gactgaaggg	tgatttctag	tcttgtcctt	gtcccctacc
44161	tactgtaaat	ttgcatggtc	agggtggaat	tcaacagaag	tgttgtaagg	taaagactct
44221	gccactcaca	aggaatttcc	ctgtgagcaa	ctcctctggg	aggccaccta	gggagatatg
44281	tgggcttctg	tctttgcagc	tgtttaggaa	cagaaggaag	gcagtttttg	cgtgactcag
44341	ttcacaagct	taactttttg	ttgttgttgt	tggtgttgag	acagggtctc	gctctgtcgt
44401	ctaggctgga	caacagagaa	tgatctccac	tcactgcaac	ctccacttcc	gggttcaaga
44461	gatttctcct	ccttagcctc	ccaagtagct	gggattacag	gcgtgcacca	ccacgcctgg
44521	ctaatttttg	tatttttagt	agagacaggg	tttcaccatg	ttgaccaggg	tggctctgaa
44581	tgccagacct	cagatgatct	gctggccttg	gcctcctgag	atgggcatag	tgctagtatt
44641	acatctgtga	gccactgcgc	cgggccctaa	gcttaacttt	tcttcttggc	agagtttagg
44701	gtcccagagt	gttattttcc	ttccaacaga	gacatccaca	gctttcacga	tggttaccga
44761	tcatatgtgc	cattcatcta	agaccaaggt	gaccacctgg	tcccaaagcc	tgcacctcgg
44821	attcttactc	gattcccgcc	tgtctcttcc	caggccctca	gttagtttcc	acaacatggg
44881	agggttctct	ctccgtggtc	tcgcccattc	gcgttcccac	ccaccctccc	cggatgtgac
44941	gacctggggc	tcccactccc	cccctagaga	gaccatccac	ttccaccctt	tcacttaggc
45001	ttcctggcgc	agcctctatc	caggaaactc	gtcccatagc	tgtactgccc	aggctcagct
45061	ttctaacccc	aagatgccac	tccgttcaga	cgttcccac	ctccatccct	gagacccggg
45121	gcaggatcgc	agcctagggg	cctcacRggc	cgctgctac	cctgttagtt	gcaacacggg
45181	gcggggcgtg	gctttgtgca	cgtcggtctc	cgggaagagc	tttacgatac	attgaccgac
45241	attttacgac	aggcggtggt	gtttgtgtgc	tgctagcttt	cYccgtgggt	tgagttgtgtg
45301	gctgcatttt	tatctctggt	ggctctgcta	cggcggtgca	gaaatgaggc	agaagtcgga
45361	agggtcgaaa	ggggaaggag	atgggggaaa	ggggtggtcc	gaaaggggcg	aacgcccagg
45421	caatcaaagt	ctgaaccgaa	cttttaccgc	gagaatccgc	taccagtc	agtcgccccg
45481	ccacctaggg	tcagtgtgtc	cattctgggt	cccagacctt	ctcgtttttc	ctgttttgct
45541	ttttgaagca	ccctaccctt	cctcttctcc	tctttggcag	tcacgtgggt	tcttgtttga
45601	atggcaggga	aaccattatt	ccaatatatg	ccctccgagg	agttaacgtc	gatttaacgg
45661	gttgtaggac	ttttcatttt	ttaagatttt	tttcacgaca	catttgacgt	gttagagagt
45721	ttcttttaaa	gccttatttt	aagatattaa	aaaaacctta	attatccttg	gattcaggtt
45781	aagagttgtg	catgcaatta	ttcccatttt	tattacatca	tgagttagca	atgaacaatt

45841	ccaccttttg	tttttcaaca	ggacggatcc	aaatttgaaa	tattctctca	caggttctgt
45901	tcctatttct	gtatctcaca	ggttctgttc	ctatttcatt	tgaccagctg	gtctaatacat
45961	ggtaattgtt	gtcagaataa	tagctaccat	ttattgagct	tgtattttga	agtcggatc
46021	atttccattt	tacagatgaa	gaagctgaga	ttcaaagagt	taagtaactt	tcccagacca
46081	aaattatact	tagatatgga	gctggctggg	caggctgtat	ctgtgacctc	aaaaaacata
46141	aaatgaaaaa	taaatacatc	gttaatcMct	atgcttcttg	acattaaactg	ttatttttct
46201	ttcttccaag	caacaattct	aagtgtccaa	gttaccttag	attaagtctt	gtcataattg
46261	catgccctaa	attaatttgt	tatctcactc	taacttgctc	cttctcaatc	taccaggatt
46321	cttttagYtt	cttctgtcac	caagtcatgt	tgattcttac	ttagaaatta	cttcttgata
46381	tagtttatct	ctcttcattc	tggccgcaa	cctcctaacc	caccttcac	actaactaat
46441	atttggaatg	tatttcagct	ttcatataca	ttatttattt	gagtcttaga	atggttttgt
46501	gaaatttgga	tgtattaccc	caattttaca	gacaaggaaa	tgggcaactca	aaagtgtgtg
46561	atcttggcca	ggcgcggtgg	ctcatgccgg	taatcccagc	actttgggag	gccaaggcgg
46621	gtggatcact	tgagggtcagg	aattggagac	cagcctgacc	aacgtggtga	aaccctgtct
46681	ctactaaaaa	cacaaaaatt	agccggcggt	ggtggcaggt	gcctgtaatc	taatcccagc
46741	tacttgggag	gctgaggctg	aggctgaggc	agaggcggga	gaatcacttg	aacctgggag
46801	gtggaggttg	cagtgaagcg	agatcgaccc	actgcactcc	agcctgggtg	accgagcaag
46861	actctgtctc	aaaaagaaaa	agaaaaagaa	aaaaaatgtg	cgatcttttc	cgttaccaca
46921	cagccagtag	gcattgggac	tggatcctgt	cttagatttt	gtgggttttt	taaaacaatg
46981	ttgccttttc	agagccattt	gagtagtata	catattatat	gtattctttt	tcactttctg
47041	aatcattctt	tgtaccgata	atattctttac	cttccctaaa	cgttactttc	aactttttat
47101	ttccaaagtt	gagaatatgc	aattcttcac	tgtctgttgc	attccatttc	aatgcctttg
47161	cctagtggcc	agggtccctt	gtaatctggt	ctttattaca	ccctgttaga	aagctggatt
47221	ttagtttagc	tgtgtctgtt	cacacaaaga	taggtaattc	ctatgcctga	cattatttct
47281	tctggattgc	cttctgtttt	gcattgtccaa	atcttgccct	atctttaaag	atccctgtgc
47341	catttccctc	atgaagtctt	ttcctactca	tttactcttc	cagttcttga	tacactcaag
47401	agctatacaa	gtggttttcca	aatcttatct	gttcttatca	gccgatcccc	ttttaccctc
47461	caaataaatt	cttacacaga	acctaataat	ataaacagct	gaaagcagag	ctgctgtggg
47521	ggaaataaag	agacctcctc	ctttttttcc	tccatgcctt	attttccctc	tcccttctcc
47581	cttctctgc	catgaccacc	accagtccca	cctctgttgc	atggcccctg	aggcaccttt
47641	gtagaactcg	ggaagtgcct	gtttacattg	tttgaaaagc	attttagcctt	tgtattataga
47701	tggactttat	ttctctcaaa	aattgttgtg	tatatgttag	ccttgccctca	ttaaattaga
47761	tgcagtgtgt	tacagtgaag	aaagcatgga	agttgtgtgt	aaaagagctg	gatttgagtc
47821	ttggctctgc	tgtttacgat	agcttggtta	cttgcccttag	cctcattttt	cacatctata
47881	aatgggata	acagtaccta	ccctccgtgt	tactgtaagg	gttcaaatga	gataaagtaa
47941	atgtgtgggc	tccataattg	ccaggccctt	ccatccacgc	tttcattatg	ttttagcttt
48001	ctctgcagtg	tgagtccctc	gccccctggg	ccacggacca	gtaccgggtc	atggcctgtt
48061	aggagctgag	ccagacagaa	gaaggtgagc	agcaggcgag	caagtgaagc	ttcatcagta
48121	tttgacagcca	cttcccatca	ctcatgttac	tgcctaagct	ctgcctcctg	tcacatcagc
48181	agtggcatta	gattctcaaa	ggagtctgaa	ccctgttgtg	aaccgtgcat	gtggtgggtc
48241	taggttgtgt	gctccttatg	agaatctaat	gcctgatgat	ctgtcactgt	ctccatcac
48301	ccccagatgg	gagtgtctag	ttgcaggaaa	acaagctcag	ggttcccaact	gattctatgg
48361	gaacagaatt	cgttatgggt	aattgtataa	ttatttcatt	acataataca	atgcagtaat
48421	aaaaataaag	tgcacaataa	atgtaattga	cttgaatcat	cctaaaacca	ccctgctccc
48481	ctggtctgtg	gaaaaatcat	cttccataaa	accagttcct	tttatcaaaa	agggtgggaa
48541	ccactgctct	acaggtgaga	agcttgcttg	ctgacagctt	ccaagctgat	atcccacagc
48601	ttttgtccct	ggaaaaagac	tggcattttt	taccttgaag	aaaagagctc	agattggcct
48661	ggcttagaga	tatgcccatc	cttgaaccaa	tccttgaact	gtgcttgagt	ggatggcgtc
48721	atacaagaac	ctgtcagctt	ctgctgttgc	ctttggatga	gaccgagagc	cccagttgcc
48781	agaagttcat	gatggtggac	agacaatatt	ttagaaatgc	attaagattc	ttaggcagat
48841	acctgggttt	gagtttcaac	tttactcctt	gatagcatct	tcattgtaaca	ggcacatatt
48901	cttgttctgt	ctacctctgt	tctgagaatt	acctgaattc	atgcattctaa	atgttcttta
48961	tgaactctaa	agtattatgc	acatattagt	aattacacct	ctggagataa	agggaagaga
49021	attacttttg	tcctttcctg	tgtgatataa	ttctttctcc	tcacattgtt	tctactYctg
49081	tatcctttcc	tgtattttatt	tttaataatc	tcataagata	tgtgcttcta	aagggaagctt
49141	ttcccatata	tatgtaaagc	cacgtttgtt	ttaaagatac	aaaagatatg	atcaaattgt
49201	aaacagtaga	gatttagcat	cttctgaact	tgagctatcc	atttggcata	tttgtatggg
49261	ttcgatgga	aaactgctac	taaatcagat	tttaaaatct	tgattgaatg	atagcatttt
49321	tgacattgtt	gaatattaac	tttccaggag	atctcagccc	tgctgagctg	atgatgctga
49381	ctataggaga	tgttattaaa	caactgattg	aagcccacga	gcaggggaaa	gacatcgatc
49441	taaaataagta	agtggatata	aagagagagc	aagcttggtc	ttaggtagca	gatctctttt
49501	atgaaattat	ctagtactac	ttgttttcca	gctcaagttt	tatttttagtt	gagaatttta
49561	gctttctgcc	ttgccttgct	attcatttct	tttttttttt	ttgcaagttc	gcctctggg
49621	gttaacgcca	ttctcctgcc	tcagcctcct	gagtagctgg	gactacaggc	gcccgcacc

```

49681 acgcctggct aatTTTTtGt atttttagta gagacggggt ttcaccgtgt tagccaggac
49741 ggtctcgatc tcctgacctc gtgatctgcc cacctcggcc tcccaaagtG ctgggattac
49801 aggcgtgagt caccgcaccc ggccgtcatt catttcttga tggcaaaatt gtctcagctt
49861 aagaatacta atagcttcta acatttattg ggcacttact gtattccaga cactgttgct
49921 aatgcattgt atgtattagc tcagttaatc ccttcaaaat tctgtgagat aggtactgat
49981 accagccaca cttgacaggt gggtaatctg agtcagagag cagttaaata atttgtccag
50041 gggtaagtaa ctagttaatt gcagagctgg gatttggaat taggtcagcc gtctccagag
50101 cccatgttct tctgtgctgc atcgctgcac atcctacttt ccttccctac tgtctcagct
50161 gaagcagcag ctttctgtca ggaaaacct ccttttgggg cccctgatct aatccaacct
50221 tcccctgctg tagaacctct ttgacacatt catccacttc ctctccataa tcattttcag
50281 cttgactttt cctacagtct caaataagag ttgtctttca tctctgcac tcttctttt
50341 ttcaaactat aatatatatt cactctatgt agatcttccc taatgaaagg gttctctccc
50401 caaccccagg tcgtctaata atcacatata ggaatgacct tttcttagac tctcttctct
50461 gatacttgag atgctgcaat attttacccc ttccttgaat cctcatgctt agtaaataatt
50521 tgctacagga ataagtaaat atcctctttt gactataaag ttattgtatt ttactgattt
50581 ttctaccgtg tctttagaag ttttcttttc ttttttgtt ggttataacc ctgattatac
50641 aaatctttaa tttctagtga acatttctac atcttatata tatttgttgt tctctttaat
50701 ttagcacatc taaagctaaa gctgccatcg tccttaaaac tcgcccctcc tcagatgtca
50761 ccagcctatt actggaaggt ctctcagctc ggcttcaacc tgcttttcca agtttttaaa
50821 cctctgtacc tttacattga aaatcttctg ctttctgtca agttggctctg cttgtatttc
50881 ttaaaatata tgacttctct ttttgcac tgcacataaa tcagcctgtc tgcctctcat
50941 gtcacttttg tgtctttagt gactgagacc tactctttct tctgtgtctc tttccatgaa
51001 tatctcccca gcacacatca gattcctccc tgaactctgg aagcattcat tgttggtatg
51061 tgtgtgtgtt ttatgacata tgttttaatc atttatacat taattttctt ggtacgcctg
51121 tctgcaaaaca tgtaaacata gaacatgaaa ggactgtgcc ttgcacattt gaacatagca
51181 ggtatttcagg cagcctcttg actgtgttag gtcctcagtt aatatttgtt gacaaatag
51241 tgagcacttc acagatatatt ctagatggat taaagttaga agacagggtg actgttaaga
51301 gtttggtctg gaggcacaaa gaaaagctgg taaaagttt ttttaaaact tcaaaaatat
51361 gtactttatt tcttatattg ttctgcatag ccctccttga tttttttttt aattgttgtt
51421 gttgttttgt ctaagacagt agttgtcaac caggggtgat ttcctcctca ggacatttg
51481 caatgtctag ggacattttt gagtgttaca acaggaagac aagaagaagg ctggtagtag
51541 atagagacca aggatgctgc taaacatccc tcaattcgct ggccacctgc cacagcaaa
51601 aactgtgtgg tcccaaacat cagtagtgct gaggttgaga aaggtggagt cagaatagtc
51661 ctttgaagct gcagcactat tctgaactag atctatattg tgtcatcgct cagggaccct
51721 tctccctcca tagtctctg gtttctttc tggtagtcaa ggggtgggaag tagtcattca
51781 tctgtgaatt ttattcttct atttccctga tagtcacaga tcttaaaact atcctgttca
51841 cgtggtgaag ggattggaga gtgacagttg ttaatgatgt tgtttataga tcatcctctt
51901 ggtagcttgt ccttaaataa ccgtRatctt gataatgtga gatgctttac tttcagggtg
51961 aaaaccaaga cagctgccaa atatggcctt tctgcccagc cccgcctggt ggatatcatt
52021 gctgccgtcc ctctcagta tcgcaaggct ttgatgcccc agttaaaggc gaaacccatc
52081 agaactgcta gtggggtgag tgattcgact catgaggtat cgacacactg ggtatctgtt
52141 Ytggcagaag tccctgctcc atgtgacRcc cgtgtagtga ggtagagggt ggggattctg
52201 aactaatgaa gtccctgtat taggatattg gctggagtgg agacctgtgc tgaatgcaag
52261 gagagcagag agagaaaaaa atataattgt gattaaagga gtagatgatg ttttgacctc
52321 tttaaaaata tagatttagg cttggtgaca atcataacag gataaatgtt ggccatctta
52381 agctggtagc tttcaatata attgatgctt atgaaaggta cttttgaaag gtagttaatg
52441 gtgcaaagta gttctgtata tggtgattat gagcactgtg gaagtccaga caacagcgat
52501 gtccctgcag caaaaggggc tggggaaggc ccagtagaaa gcacacagct tgcttgacct
52561 ataagccaag aagagcaggc tctaagtgtc atgggagcaa ggacacaggg gctaaggcac
52621 agaggtagaa atgacaggaa ggcgtgggtg cacgcaggaa atagcgagta cgccagactg
52681 aagggtgtaa acagttcagg agagtaggac tgtgcgttga acaggtagtt tgggatcagc
52741 ttgtaaaaat cttaagtgtc tggcaaaagca gctagataat aggaaacaaa agttttgaaa
52801 aaggtagtaa tatgaacaaa gttagattttc agaaaattaa tttggatata acatgcaggg
52861 tagtttgagg acaaaaagac tagcttagat gctcttggtg tgatgtcagt atgcagtgtt
52921 aaggacttga gttaggatgg agactctagg atgaaaaaag atgagtgtga gatacatgca
52981 cagaatttga tagaattaa accctcctac catatggctc ccagcaccaa attctgttcc
53041 taatgtgatc ccacaagttc attctgttat ttgaattggg atattatatt ccaggcagct
53101 tgtgttgact tggtgacct atctcctgga tcaccttacg tttgtggcta attaaacctt
53161 aaaattttta catacaatgt accaccaaac ctcatatcct cttctttgta cttaagtgtt
53221 tttgttttgt tttgttttgt tgtatctagt gtaaaaagct tgttagccag gcacgggtggc
53281 tcagccctgt aatccagca ctttgggagg ccgaggtggg ccgatcaciaa ggtcaggtca
53341 agaccatcct ggctaacacg acgaaagccc atctctacta aaaatacaaa aaaaaaaat
53401 tagccgggag ttgtggcggg ggctgtagt cccagctact ccagaggctg aggcaggaga
53461 atggtgtgaa cccgggaggt ggagcttgca gtgagctgag atcatgccac tgcactccag

```

53521	cctgggagaa	aaagcgagac	tccatctcaa	aacaaaaaac	aaaaaaactt	gttaagaaaa
53581	actaatagtc	catgcccctc	acctcccttt	ttctacccta	gggcaaccat	ttttaactct
53641	tagccaatth	cttttagcatt	aacttccata	tccataaata	aaataacatt	ctttacataa
53701	tagataagtc	ttgactttct	tttttttttt	ttttacttga	gacagtcttg	ttctgttgcc
53761	caggcaggag	tacagtggta	cgatcttggc	tacttgcaac	ttctgectcc	caggttcaag
53821	cgattcttgt	gcctcagcct	cccaaggagc	tgggattaca	ggcatgtgcc	acaatgccc
53881	gctaattttt	gtattttcag	tagagacagg	gtttcactat	gttgccagg	ctggtctcga
53941	actcttgacc	tgaggtgatc	tgccgcctc	agcctccaa	agtgtggga	ttacagacgt
54001	gagccactgt	gcctggccga	cttttttagag	tttagcatta	tgtgtgggct	tgccattaaa
54061	gaagacagaa	acttagcaac	ctttcagcct	gactggcaaa	ccgaggcttc	tgtgatacca
54121	ccctctctgt	ttcttctct	gtctctgatc	atctctgtcc	tatcttggt	accttattgt
54181	gcttacctct	aaaagtgtat	gttttctgtt	gtcttgggt	cattgtgttt	ggtattttac
54241	atgctttctt	tggtaacctc	atccatttga	tgattttagt	attgatgtat	gctgactccc
54301	agcatgaacc	attccctgag	cttcagactc	ctgtcagatt	gtcagttagg	catctgtcct
54361	ttgactgcct	gagaacctcc	tgaagtatag	cataaccaa	actaatacca	gacttgccta
54421	ttcacctttc	ctgtccatgt	tagttcatgg	taccacgtg	cactcagttt	ccaaaaaatg
54481	tgaaatgcat	tttccgttcc	tactgccact	ggtaagaatg	ctttggcctt	tattatttct
54541	tgtttagatt	atttcagtca	tttcttacg	catctatth	gtcctaccta	catgaaatgc
54601	atctttaaca	gtgtcacaa	agtgtatct	tgaaagtatc	agaaacacac	agctggtggc
54661	cattccctat	ctactgccat	tctcaccttc	atgcttttat	tcttgcctgt	ttctgtatgg
54721	tgctttgttt	atthttata	gtagtthtag	gtttatagag	aaaaaaat	tatacacttc
54781	ctctagcacc	ttcaccccta	ccccagtttc	cctattatta	atatcttgca	ttgtttggta
54841	catttggttag	aattgatgaa	ccaatattgt	tgcatthta	ttaccaaag	cctgtacata
54901	cattagaata	cactctgtgt	tctacattct	gtaggthttg	ccaatgcata	atgtcatgta
54961	tccactatta	ctgtgtcata	tgaatagtt	tactaccct	aaaatctcct	atthgtggcc
55021	ggcgcgagtg	gttatgcct	gtaatcttag	cactttggga	ggctgaggca	ggcagatcag
55081	ctgaagccag	gagtttgaga	ccagcctggc	caacatgggtg	aaacctgtc	tctactaaaa
55141	atacaaaaat	tagccgggca	tggtggcggtg	cgctgtaat	cccagctact	tggtgaggctg
55201	aggcaggaga	atagtttgaa	cccaggaggc	agaggttgcg	ttgagccgag	ataatgccat
55261	tgcactccag	cctgggtgac	agagcgagat	tctgtctcaa	aaaaacaaac	aaacaaacaa
55321	acaaacaaaa	acagctccta	ttgtccctt	tccctctgca	tgthctagac	gtaacctgac
55381	ttccactgat	tgthttattg	tctthtaata	agthtgcctt	ttccagagt	tcatgtacag
55441	taattggaat	catacagcct	ttccacttag	caatagcat	tgaagtctgc	catgtcttht
55501	tgtgacttgg	tagctcattc	ttthttttta	attactgaa	gataatccat	tgtacggatg
55561	taccactatt	tgcttattca	ttcacctata	gaaggacatc	ttggttgctt	ccaattthtg
55621	gcagthttta	acaaagctct	gtgaaggtht	ttgtgtccac	ctacatttht	agcttacttg
55681	agtaactgtc	acaaagtgtc	actggtagat	catatagtaa	gactatgttt	cactthgtta
55741	aaaactgcaa	actcttccag	catggctgtc	ccattthtgc	ttcccaccag	cagtgtgtga
55801	gcactctgat	gttccacatc	cttgcttaaca	cttgagagt	tcggtgttht	ggattthtatt
55861	taattaattt	atthatttht	agacaggtt	ttgtctgtc	actgaggcta	gagtgtgtgtg
55921	gcatgatcac	agctcactgc	agcaacctcc	caggttcaag	ctatcctccc	acctcagcct
55981	cccaagtaac	tggtgacaac	ggcatgcacc	accacaccag	ctaattthttg	tgthtttht
56041	agagacaggg	tttcaccatg	ttacctagga	tggtcttgag	ctcctaggct	caagggatcc
56101	tcccagcttg	gtctcccaaa	gtgtgtgtat	tataggcgtg	agctatgggtg	cccagccagt
56161	gtthttggatt	ttagccattc	tcacatttga	acagtgttat	ctccttgtht	tttagtht
56221	aattccctaa	tgacatgatg	ttgagcatct	ttccgtatac	ttattthtcca	ctgtatatct
56281	tctthattga	gatattctagt	cacatcttht	gccctgttct	taattgggtg	ttthcttact
56341	aggtthtaag	agthtcttht	atatcttagc	tggaagtgggt	ttatcaggt	ccaattatgc
56401	agatatttht	tcctagtctg	tggttgtct	tttcattgtc	ttctcagag	cacaacttht
56461	aaatatagac	aattaggtcc	ataacccatt	ctgagtht	atthtggtht	gtgtht
56521	tcttgaaaca	ccatggtht	taacttggat	acattactgt	catctaaagc	tcagacctca
56581	cttaagtht	accagccgtt	tcaataacat	cccacagaa	ctagthcaga	atcacctgtt
56641	gcattthatt	gtcatatct	tttagtctg	acatttccct	tgtcttht	ggactccgtt
56701	atcttaacgc	ttthtgaagat	ttctggcaag	ttattthtga	gcaggtccct	cagtgtgggt
56761	tcactagctg	ttthtctatc	atgagattca	ggthacgctg	ctthtggccg	tgctcatag
56821	aagcagcact	acgttcttct	cgctcatctc	catccagtg	tgcgaggtt	tggtthtct
56881	atcactgatg	ttctcattt	tgatcaagg	gctgtccacc	agacttacc	tctgtcaagt
56941	tattthtttc	cacttht	taagaagtgt	tgatggaga	aatactgaga	aactaggtg
57001	atatcctgtt	tctcatcacg	taccagtht	actcttht	ttgtgtgaag	gaattaatgg
57061	ttctctattt	gggtggtht	catctgttac	tatttht	gatgcacaa	ttatttggga
57121	cttgaccagt	gggagcctt	tcaagctgat	ttctatgtct	ttthaaatg	tcctcatcat
57181	tctthtga	gtthtctagct	ttctagcaca	ataaatgtt	ccaggttgg	ccagacatgg
57241	tggtctcagc	ctgtaatccc	agcacttht	aaggacgag	tggtgtgatc	acctgaggtc
57301	aggagthtga	gatcagcctg	gccaacatgg	caaaacctg	tctctactaa	aaaaaaaaa

57361	tacaaaaaat	tagccaggca	tgggtggcaca	tgcctgtaat	cccagctact	tgggaggctg
57421	aggcaggaga	attgcttgaa	cccaggagggt	ggaggttgca	gtgagccaag	atcacgccat
57481	tgcactccag	cctgggcaac	acagtggagac	tccgtctcaa	aaaggaaaaa	aaaggaaaat
57541	gttccaggct	tatcttagac	tttctttgct	ccagccctgg	aatcagccat	ttccccaagg
57601	agccctgggt	tcttttagtt	ggagaaggat	atttagatac	taagacctgg	gtcctagggtg
57661	tgcttactgc	tgttagggtg	ttgctgctgc	cagactctct	cagtggacca	agcgaggaca
57721	tatatatgta	tagctgcata	cataacatgc	acacatacat	gtaacacatt	tccatttgta
57781	tttattttatc	agtctaccat	atgttgaaaca	ctctgattgg	ccacaatacc	tttaattcca
57841	cccagcccac	agagttcatt	ctgcttctct	ctctttccat	gtttatagct	acttctctga
57901	tagtaagaag	cctggcttgc	tttcactttt	gtaaatggcc	agatttgacc	aagtgccttg
57961	gatgtaacca	atcttgcgtc	tctgccactg	cctcctgtcg	tcacctcact	gaggctctgt
58021	cagacccctc	tgaggttatt	tacacccaga	ccctgaaaca	tgaagctgct	agtttaatag
58081	tacctgctgc	aaatattgag	atccagtgtg	ttcatgaggc	gtttgagtca	caaagggttag
58141	gtttatatat	aatttcatag	aattgcttaa	agaaattttt	ttcttacagg	ctgtttacta
58201	agacaatcag	agagagaaa	actaagaatc	actttggctt	taacagttaa	tttgttattt
58261	tgtacttaac	ttattgtaaa	atggaatata	acttcacata	tatattacat	acggacaatt
58321	taaagatgat	taatatgaa	cagagatcac	tcttgtaacc	attgcccagc	ttaagaaata
58381	cagcctcggc	cgggcgcggt	ggctcacgcc	tgtaatccca	gcactttggg	aggccgaggc
58441	gggcggatca	cgaggtcagg	agatcgagac	catcccggct	aaaacgggtg	aaccccgctc
58501	ctactaaaaa	tacaaaaaat	tagccggcgg	tagtgccggg	cgctgtagt	cccagctcag
58561	tgggaggctg	aggcaggaga	atggcgtgaa	cccgggaggc	ggagcttgca	gtgtagccag
58621	atcccggcac	tgcactccag	cctgggcgac	agagcgagac	tccgtctcaa	aaaaaaaaaa
58681	aaaaaaaaaa	aaaaaagaaa	tacagcctcg	tcaatacctt	tgaagcccct	ttttgccact
58741	ctctggttgc	attcctctcc	ttccttccga	gggataagca	ctctgtggag	ttttatatta
58801	attatcctat	aatagctttt	ttattcatct	ttcttttata	gattgctgtc	gtggctgtga
58861	tgtgcaaac	ccacagatgt	ccacacatca	gttttacagg	aaatatatgt	gtgtaagtat
58921	ggtgatttta	ttaaattgta	tgtatgtttt	aattaagcta	aatatgcccc	ctctagccct
58981	tagtcagtac	atcctggtaa	tgtttaaaac	ttcagcttaa	tagatttata	gattactcct
59041	ttcaaacaag	caaccattgg	tagatatttt	agtgccttaa	aattggaata	tataaggccg
59101	ggcaaagtgg	cttacgccta	taatcccagc	attttgggag	gctgaggcag	gtggatcacc
59161	tgaggtcagg	agttaaagac	agcctggcca	gcactgtgag	actccgtctc	tactaaaaat
59221	acaaaaatta	gccgagcgtg	gtggcatgcg	cttctagtcc	cagctactta	ggaggctgag
59281	gcaggagaat	cgtttaaaac	tgggaggtgg	aggttgagct	gagccgagat	cacgccactg
59341	cactccagcc	tgggcaacag	agtggagactc	catatcaaaa	taaaataata	aaattagaac
59401	atatgaatat	tttaatttat	tgcaatatac	aattctaaaa	atgtaggtta	tggaaactac
59461	aacagtagac	attgggatat	gcaactcaaa	acagcacatt	ctgttaaaact	cataaattgaa
59521	acatgggaat	atgagctgcc	taaattccac	gtcagaaaat	taaaatgaat	ttggatcaga
59581	aacatatcaa	aataaaaaat	tatccttata	cgtaaccctt	agatttctca	aactcaccta
59641	tttgaaagaa	ttagcggaag	agtttgcatg	attctgggta	ggaagataaa	aggggggaaa
59701	gattttcttg	agtgtggttc	caaagagaac	actgagtatt	gattcaaaact	taggaaaaact
59761	tcttttctgt	tatttgcccc	taaaagctta	aacctctgaa	ataaacacaa	ctgcagttat
59821	tttgaaaatg	ggtgtaataa	tgtcccttta	catattttatc	ttattttactt	cttggttagaa
59881	aatacattat	ttacagatac	tgcaaaaggta	gattcttctg	ttaggatttg	aaaaggaagc
59941	tcatattgta	tcaggattct	ttggagatgg	tagatgcttg	gaactagctg	attgaaactca
60001	gttttgatc	tgacattctt	gttcttttgt	tgtactggca	gatactgccc	tgggtggacct
60061	gattctgatt	ttgagtatte	caccagctct	tacactggct	atgaggtaca	gtaactttga
60121	ggctgtcctg	atgaaatggt	gcatcatgct	ttacctgtag	tatggtttta	ccagtagctg
60181	ctttctgaca	atttttttgt	tttggttttg	ttttttctg	atttttaaaag	ctgttcattc
60241	accaaatatt	tgccagtgtc	tactagtgcc	ctatattatt	ctagccacta	gagaaatact
60301	tacataagca	taaataatac	ttacttctta	tcccaaatgc	tggtctaaag	gctttacaac
60361	tgtaaacctt	attgcagtgt	tttgaagata	gacactatca	tgatcccagt	ttgccaatca
60421	ggaaactag	gcatagcagg	cttaaacagc	tggcccaaag	acacactatt	agtaagtga
60481	aaccagggtt	cgaagtctgt	gctgcttacc	actacattgt	actgtcactt	tagcagtgg
60541	aaatggacag	ggtcccatgt	tcttctggaa	tttacatttt	tttttctttt	tttttttttt
60601	ttttgagatg	gagtcctact	ctgttaccca	ggctagagtg	cagtggcacg	atctccttca
60661	ctgcaacctc	cacctcctgg	gttcaagtga	ttctcctgcc	tcagcctccc	gagtagctgg
60721	gactacaggc	aagcaccacc	acgcccagct	aatttttgta	tttttagtag	agacaagggt
60781	tcatcatggt	ggccagcctg	gtctcgaact	gctgacctta	agtgatccgc	ctgccttggc
60841	ttcccaaaat	gYtgggatta	caggcggttag	ccactgtacc	cagcctggaa	tttatatttt
60901	ataatggaa	ggtagacagt	aaagaaacaa	gaaaaagtat	caggcaactca	aaatctccta
60961	agcagagatg	taaaatcaag	taagggtgatg	gcagctgaga	tgggttagct	gtggctgcta
61021	ggaagggatt	ctgtgaagtg	agattgaagt	tgagtctgaa	tgacaagaag	gacctagtcc
61081	taagaatatg	tgaaaggggc	attcctagca	gagaagtcac	tagaatgagg	cccaaggaag
61141	gaaagacatg	gggggtgtgt	gcagtgcagg	gggtgagggg	agcatgaggc	tgggagtgc

61201	aatgagctgg	ggtttgtaag	caaaggtaag	gagtttccat	tttagcgtag	gagtcgatgag
61261	aagctattag	gatttaaggc	aggggaatga	tacaatccaa	tttaggtttt	ttgaaagatc
61321	attttgatgg	ccatgtagag	aaagggttag	agtggagacc	agaaagaagg	cagaaggcca
61381	gtgaggtgct	tttgaaggag	tcaccctcct	cagaacacct	cagaagccag	gaaagcttgt
61441	gacttttttc	tcatatctgt	tttcattttt	tttcttggtt	tctagggttc	aaatttttta
61501	aaatacaaga	ggaatgattt	cgtgaaaagg	cttcctctca	ttcctatccc	ctagccactc
61561	tttttcactc	ttccccctac	cagttagtgt	tgtttggtat	ttcctcccca	tccttcaggy
61621	atatttttag	gaacacagca	tacacagggt	gcctatccct	tatgtgaaat	gcctgggacc
61681	agaagtgttt	cggatttttg	atttttctgg	attttggaat	atttgcatat	acatagttag
61741	aaatcttgag	agtggaaccc	gaatctgaac	atgagattca	ttttgggttt	atgtacacct
61801	gatacaccta	ggctgaatta	attttataca	atatttttaa	taattttgtg	cgtgaaacaa
61861	agtttggtgt	aagtactatg	tgtggaattt	tccagttatg	gcatcatgtt	ggcactcaaa
61921	aaattatgag	gtttggagca	tttttgattt	ttggattagg	gatgctcaac	ctggacgtat
61981	tctttatttg	actgattcca	tataaggtag	catatcagag	tcctctttca	ctttgccttt
62041	tattttacaa	tatcttagtg	aacattttat	gtcagtacat	tgtttcgtgg	ctgtagaagta
62101	ttccactgta	tggtggtagc	cattacaac	tgtgtggtat	tttagaagct	tacttaaaag
62161	ttactttatt	atatgattct	ggtatatgta	tgcacatcta	tgtctgtgag	cagaacactt
62221	tggtgaccct	gggattccag	aagtgtttat	acaaaagaca	gatgtgatcc	aaggagacac
62281	cctgctgttg	aggtgtttat	gacagcgtga	gtggacacct	gccagatgcy	attcaggaca
62341	ttattttgaa	ccctgacaag	actgagaaaa	attaatgcgg	gtacaagcca	cgttttcagt
62401	gttcggaacc	atggagagtt	ttttttaaaa	tacagtcctt	ttgaaactac	tttttagttt
62461	taattcaRtg	tgggcataac	aatatttttc	tcttctagcc	aacctccatg	agagctatcc
62521	gtgccagata	tgaccctttc	ctacagacaa	gacaccgaat	agaacaggta	catttttaaa
62581	aaacatgttt	cttaaaaaatt	aggtgtttat	acttagtaag	aagccattgt	tgcttgattc
62641	aaattgaacc	tgaataaaga	atgaaaaagg	tgtttttccct	ctttgtaagt	tttcaatac
62701	caatttgagg	agggagaatt	tgccatgcct	agcaagggtca	aaaacactac	ttcttttaaa
62761	gactgtattt	attgttttaag	ggttttatat	tctctaagtt	tttttgaatt	tgtagaaagt
62821	catttgtagt	atgaaatttg	tggaataaag	atgtatgaaa	gttcttagac	aatgggtggg
62881	tgtgttgact	tttaattttcc	aaaagtcaga	ttaagaagta	ttttgactgg	ccatgcycaa
62941	tggtctcatgc	ctgtaatccc	aacccttttg	aaggctgtga	cagcagggtca	cttgagccta
63001	agagttcagc	accagcctgg	gcaacatagc	aaaaccccat	ctacaaaaaa	tacaaaaatt
63061	agcatggcat	tgtggtgtgc	acctgaagtc	ttagctactt	ggaaggctga	ggtggaagga
63121	tcccttgagc	ctggaagggtc	aaggctgcag	tgagctgtga	tcataccact	gcacttcagc
63181	ctgggtgaca	cagcaagaac	ctgtctcaaa	aagaagtatt	gtgacagatt	tgttgggtgg
63241	aaataggaaa	tttccctacaa	aggagtacaa	agaactagtc	ggggtatggc	attgttctct
63301	atcatgtgca	tgggtggtgt	tacataactc	taaatatcta	tcagctctca	tccattgtac
63361	acttaaagtt	agtggatttt	atcgtattta	aattatacct	cagtatgggt	gactaaaaac
63421	aagtactatg	tacatgacct	tgcagtgttc	aagaaatctg	aacattaata	cagattttcct
63481	ttattttacaa	gtttattttta	aacttgtcca	atttaaaaaa	tgtaaagcac	tgtccatagt
63541	tgtaatagta	atgtatagta	ggcacatcca	agtctaaagt	agataatggt	acataacctat
63601	agtggaataag	tgtctctgtg	gtttgtttat	tggtctattg	gtgaataactg	ttcagtttta
63661	atatccactt	tgtctgtcacc	caagcgtatg	aggaacagga	ttgtcgggtga	caggagagga
63721	ctccatctgg	gggagcccac	atttttccaa	acagtgggtt	ctaaactgac	ctttgcttca
63781	atttcttttt	gggctatgat	agtttaattta	tttaaaatgt	aaaactattg	agcatgaaat
63841	gcttatgttt	accaaaaaaa	ggagcatagt	ttacaagatt	tagaaatgaa	catagagcag
63901	tgattctttt	cttaaatgca	ctagaattac	ctgagtaact	tctccaaaac	gtgtcacatc
63961	ttcacctcgg	gaggttctga	tttaagctgc	tttaaggctcc	tagcaagata	tttttaaaaa
64021	ctaccagggt	aattttgatg	agtatctctt	gtaaaagaacc	atagatacag	aaatagagta
64081	ttcttttagt	gttgatataat	atgtacacac	atgcataatat	atagtttttc	tgtatacgtt
64141	ttttgccatt	ttcagaaatt	agtgttaatt	tcaataccta	tttttaaaaa	ttagaatctt
64201	ggcttattgt	agtcacaacaa	atgaaagatt	tgtatcattc	tctccactag	tagaggagac
64261	ctaattttat	tattattatt	tttttttttt	taaacagagt	gtcactcttg	ttgccagggc
64321	tggagtgcag	tggcacaatc	ttgggtcact	gcaacctccg	cctcctggtt	caagcgattc
64381	tcttgccctca	gcctccagag	tagctgggat	tacggcatgt	ctggctaatt	tttatatttt
64441	tagtagagac	ggggtttcac	catgttggcc	aggctgggtc	cgaactcctg	acctcaggtg
64501	atctgccagc	ctggctctcc	taagtgtctg	gattacaggc	atcagccacc	gcacctggcc
64561	tgaatatttc	atttttaatc	agactttcac	tttttttttag	aaagcagact	tgaagtgcct
64621	cctgtgcctg	gaatcatcca	tcaatttttag	actgctgtct	tgatttttct	ttccaatcta
64681	ttcttttttt	cttccattac	atcaaatcct	tattatgtta	cataatcatt	catgtatcat
64741	ttttggccat	aaatggccac	ttttgtcct	ctaggcttac	caccaagtct	gacataaaa
64801	acatgatcaa	taaatactta	ctgttttgca	aattgtatta	tatttgcctt	tactgctttc
64861	tagtttatat	tcttcgtgtt	tttaaaatttc	cgctttgtag	gtattcaagt	caagcctctt
64921	atttgttatt	ttatattctc	attctccctc	cttatttgag	ttgtactcac	tttttctttt
64981	cagacttggg	cctatttttt	atccacagaa	ttagctaagt	gtgtttcatt	acttctgatt

65041	tttaaactgt	actgatgaaa	acactgcaaa	ataagagatt	tgcaatgcct	tcttagagta
65101	gttccttatg	cttatatcat	tctaagtctg	atgaatttgt	ctttcagtta	aaacaacttg
65161	gtcatagtg	ggataaagt	gagtttattg	tgatgggtgg	aacgtttatg	gcccttccag
65221	aagaatacag	agattatttt	attcgaaatt	tacatgatgc	cttatcagga	catacttcca
65281	acaatattta	cgaggcagtc	aagtaagaaa	ttcttatttt	atcatagtct	ccagagtggg
65341	tgtcagttta	tgctcctagc	agtagtctac	gagaatgcct	tctgccctgc	atccacattc
65401	ttacttctca	taacttttct	tgtttcatgg	gaaaggatta	tttcagtga	aataatgctt
65461	tcgctgaaat	aatcttttcc	agtgaataa	atcctttcac	tgaaataatc	cttttaaaga
65521	aaaaatgaat	acagtttggt	gacatagtag	tattctcaaa	tagagagatt	cKccaaatat
65581	gggccatgaa	tgctttttca	gcccacccaa	gaagggggct	gggagttatt	tagctgctgt
65641	aaaatcagca	aaacaaggta	ctgttattta	aaaacctaat	ttatagtaaa	tatttcataa
65701	ttattaaata	gtaactgttt	atgattggga	tcttagtctg	tggtgtgctg	ctgtaacaaa
65761	atatctgaga	ctaggtaatt	tataatgcac	agaaatgtat	tggttcacag	ctttggaggt
65821	tggggaagtcc	aatgtcaagg	tgctggcatc	tggaagagc	cttcttacta	cgatcatcaca
65881	cggcaaaaaga	caagagaaac	aaaaagcgga	ccaattcacc	cttttataat	ggcattaatc
65941	ttacccacaa	ggtcagatcg	cctctcagag	gtcccacctg	ttaatactgt	tacaatgaca
66001	atttcaacat	gagttttaga	ggggacaaac	tcatgtctgt	caaatacatgt	gttattcaaa
66061	ccacaagtta	atttagtcat	tttgaattct	agttgacaaa	attatgcac	attttgacac
66121	cttgttttta	gcaagaagaa	tactacaggt	tagtatgtag	ttcagtgtat	taagaagtga
66181	aagtcttaaa	atagtttttg	ttctcaggag	ttgcaggaac	acctggatag	ttactatttt
66241	ccttatttaa	caaatccttc	tttagggccc	actagatact	atgtgctgtt	ttagatactg
66301	gaaagtgggtg	atatttaggc	tgagaccgga	agtacaagga	ggagttaggc	caacagcaag
66361	aaagataaaa	taaaggctct	gaaatgggaa	gcagcctggg	gtgttctagg	gacagagagg
66421	aagccagcat	ggttagatcc	tgtgttaatc	catttggtgc	cctataaagg	aacacctaa
66481	actaggtaag	ttataaagaa	aagaggttta	attggttcca	gttctacagg	cttcacatga
66541	agcatatttc	tgccatttgc	ttctggttag	gcctcaggaa	gcttccaatc	atggttgaag
66601	gtgaagggga	gccagtacat	cacatgacgg	gtgcaggagg	gtgccacact	cttttaaaca
66661	acaaatctca	cgtgaaacaa	ctgagcgaga	actcacttat	caccaaggag	atggttgtaa
66721	gccattttatt	tatgaggaat	ccagccccag	gacccaaaca	cctcccacca	ggccacacct
66781	ccaatatttg	ggatcacatt	tcaacataag	atgtggaggg	gacaaacacc	caaaccatgt
66841	cgatcccg	tgaaacaag	aaagaacgct	atgagacggg	gttgaaagtgt	taggttaggg
66901	ccaagataca	tgcttttaag	gagttgaaat	tttatttgaa	atgcagtagg	gagcagatga
66961	aggggaagt	gcaagttctg	gttaacagg	agggctattc	cggctactgt	ctggctaatt
67021	gattagagga	gtaccaggg	gaaagtggga	aaccagatag	gaggccatct	gattacatcc
67081	atccctgctc	agatgggggc	ggcagcagtg	gtgatggaga	ggagattgag	atgggggtaa
67141	gagaaaagaa	gggatcaagc	ctgacactaa	ggttttggct	gtcagaaatg	gtaggaagg
67201	agggtagggc	tgtatactga	ggtgggaaag	atagcgggaa	gagcagagaa	aatctcagag
67261	ggagaaaatc	aagagttctg	ttttggatgt	gtgaagtctg	aagagcctgt	gatacatcca
67321	agtggagatg	tcgggtgggt	ctgaatgcaa	gagaagtctg	agctgacgg	acgaaactgg
67381	ggattatcag	ctcataggta	acattgacaa	ccatgtaagt	ggaggagacc	acctctagg
67441	agagttgtg	atctccaagg	tatactaagt	gaaaagcaat	ttttagagca	attcttatag
67501	tacgatccca	ttatttgtgt	gttcacgcac	atacacacac	acatatctgt	atataaatgc
67561	atagaaaagg	tggcagaata	atggtcatca	tagaccttag	agctgaggag	gaaaggacat
67621	gggaaatggc	agcaaaggag	gataatttaca	tttgctctgt	atacagtggg	ccaggtgttg
67681	tcatgggtgt	ttaatattca	cttatttagt	actcatagtt	agccttttag	ttaatgtttc
67741	agattatctc	tgttttatgg	gtgaggaagc	tgaggcacag	agagataagc	aatgttccca
67801	aagttgcaga	ggtggttgg	ggtagaatgg	gatataaatc	ccaggttagct	ttgctttcag
67861	agcctaactt	tgcaagctgt	gctaggtgtc	agaatgtgag	tgtgtctgta	tgtagtgca
67921	catgtgtgtg	cacatcatca	gagcttgaag	atcttggaag	gaatatggcc	tgtttttctc
67981	tgccctcctt	ccctaccacc	ctcaggcttt	tctctggctt	ctcttttata	tggggtgagg
68041	gtttcatata	gctaattata	aggttgttca	aatagtgcc	cctcttaaga	ttttttgtgt
68101	aggacaaaat	tttgataga	cctaagagtg	gtttttatta	ccctgtaagt	aaagcagttc
68161	ttggcacata	gtaagcacia	gtaaatgcgt	gaatgaattt	tgaatgaaca	gttagctaatt
68221	gacctgggta	gggttgccctc	ttggaattgg	gggcagccac	atctttttgt	gccctcgcta
68281	ctcccctac	ccccttaact	tcctttgttc	tccttggtt	tgtaaaaagt	aaagaagag
68341	aggagctttt	tcataaaaat	taataccaag	ggtagctcaa	agagcccatc	tgaaaggttt
68401	ggcagctggg	agagtttgtg	tggaacagcag	cccacttctg	tttgattgac	tctagggaat
68461	gcaacaggtg	aattctgtgt	ccgtgaatct	ggacctgtag	cattgtgtatt	tcttcgtctt
68521	acaggggctt	tagtaataga	ggagatggcg	actgcattgt	tactgctcgt	tcaaaactga
68581	tcaagaggcc	ggcgtagtg	gctcagcgc	gtaatcccag	ccctttggga	ggccaaggcg
68641	ggcagatagc	ttgaggccag	gagtttgaga	ccagcctgac	caacatgggt	aaacttcttc
68701	tctactaaaa	atgtaaaaaat	tagctgggca	tggaaggctgg	tgccctctagt	cccagctact
68761	tgggaggccg	aggcacagaa	acacttgaac	ccaggaggca	gaggttgca	tgagccaaga
68821	ttgcactact	gcactccagc	ctggggcgaca	gagtgagact	gtgtctcaaa	aaacaaacaa

68881	caacaaaaaa	aactgatcat	taatatgagt	catacttagt	aaatgctgaa	gtcttcaaac
68941	tttagaggag	taatgatata	atccagctaa	ttactcttaa	taatactgaa	aaatcaaaact
69001	ataccttaga	taaaatgtga	ttgaggaaaa	acaaccttta	ttagttcaaa	gccaggcgac
69061	ggggatggca	gcagaagggt	ctctcagagg	gttgctgacc	acagttcatt	cagctctgaa
69121	aattccctgg	cagggacatc	tatgaagata	agtttttctc	tgcaagctta	tatacttctg
69181	tactcatttc	ttggacctta	atatgtaagg	tcttcttata	ttgaagacct	tacatattaa
69241	gtggaattga	gctgtaaata	tcttagactt	gcctctctcc	cccataaaaa	tttgccacta
69301	agcttttcat	ctcctacagt	ttgggtcccc	tgaggatat	gaagcaggcc	aactaagatc
69361	tgcatagtga	acttttagta	tgtatctagt	ttgacatttt	catcaattga	aagtaaaaaat
69421	tttggtttat	tcttggtgta	acattttatt	tttgagaaa	tgttctagt	ctaattggtgc
69481	ttgaatgtaa	gttttccatc	attgggttga	aaatagggtt	gtctagtcca	gcgagctcag
69541	tgcatatcat	gatgtgtttg	tagaaaaagc	cctgtggaag	agaaatcctc	tttcagtaat
69601	attctaggca	gtgccagtgt	tgttttgttt	ctgttcttga	atttacctca	agagggcaac
69661	gaacacttta	ttttcagata	aaaatttata	tatgatttgg	gtcttcattg	caacacatct
69721	catgaatgcc	tcttgagaag	taatgaaagt	acaatctggg	agccataaaa	ccatccataa
69781	attacactga	attctgccaa	cacacacttt	aaatgttttg	ctcttttctc	tcagtctcta
69841	tatttttatg	agatcatctg	gaaaaaaaaa	agacctgatt	tggtgctgtg	tggtgctttg
69901	ttaaggtaaa	gttttactac	aaaccctca	taatagagtt	tgtatttgtt	ttgagggaaa
69961	ctttgtattt	gaggaaataa	tagtctagtt	tgtgctatag	aactagagac	agaaagtatt
70021	ttcaagtgtt	ggcataattg	tgaataaaaa	agcagcccag	agaagtgttg	gttttgacat
70081	aatgtggccc	tcggaaatgt	ttggatttga	ccttgccctt	ctctctcctc	ctgcccagag
70141	tctatgagt	aaaactgggt	ggtttgcaca	gcgtagccca	ctgctcttag	atgtaagggtg
70201	atgaacttca	tgttttattt	acttttggtt	ttgcttgctg	actacataga	tgtaaactga
70261	ctttcattag	cttagcaggg	ttttttaaag	attaatttta	aattaggtta	aaaatgatgt
70321	attgtgacct	atgagtcagc	aagcagcatt	taaggttaat	agtctgttca	cgttagggtc
70381	aagttttact	gctgtgttgg	ctcaggtgtc	cctgctatgt	tttcatatgt	tgaacctgat
70441	taaagttttg	cttcttaaaa	gaataggagt	taaggtaaa	aaaagcccca	gcaagcagag
70501	cctggttatt	atttatggca	agctagtagc	aagcagtggt	ttatatatat	tctcgtggat
70561	ggataaattg	gaaagttag	tgaacagaga	gttcaaggac	aaaacaggta	tggtttttgt
70621	gaaggctcat	taaatcaagc	aaagtgttaa	tcactcagta	ctatcagctg	gactgagatt
70681	cttagctagt	ctccagagag	caacaattac	tggtgactgt	catcgtgtaa	caatcaggct
70741	ctggagatga	aaaagaccgg	tagtgggatc	tgagtcaccc	attcactaga	atgcaaagt
70801	tgccaaataa	ctccaacaac	cttttaaaat	agttttattc	ctttttaatc	agctttgccc
70861	agaagcagtt	ttacattcaa	tctttaatgc	tccttggtctg	ttttcacaag	atgcaattta
70921	aagggtgatt	accatttaaa	aagttagtga	gtcatacttt	ctccctgtgg	aatttttaaa
70981	tcatttcccg	ttccttccct	ttcccccggc	ccgcccccca	cgccccatta	atgacttttag
71041	atcctccaac	tatgtttctt	cctgtctgag	aaaagctgaa	gtgctaggta	atgctaggta
71101	ccaggcccg	aagacaattt	cgtagacttg	cacagctgca	acggaagcaa	aaggaacct
71161	cagagacctg	agagttagtg	actgtggccc	tgctgccctg	ggcgctcatt	ctggcaggcc
71221	tcaggacctt	ctgcatttct	gggctttgac	gctgacactg	cttatctctc	actttttcta
71281	ttgaccattt	tactttctct	tttggtcacc	cagatttcca	tacatggtga	ccaggatcct
71341	taacattggc	cagagaacat	aggatacaat	cttagtcact	ttaagagagt	tgatatggtt
71401	tttctttcag	cattttattt	gaaacaaaaa	ttaaacagtt	tttagtgagc	atccacatac
71461	ccatcaccta	gattctacga	tacttgcttt	atcacatata	tgctcagttc	actatccatt
71521	catcagtgct	tctcgcgtgt	gcttgccgtg	tcttttttga	tgaatttcat	agtaagttgt
71581	atgcttcagt	acacttctcc	cgggataact	catcatgcat	atcactgact	agtggtcact
71641	gtctgcagtg	tttttctttt	gaagtaaatt	tacatacagt	acaaaacaac	ttgtggcggt
71701	ttatatatat	atatatatat	atatatatat	atatatatat	atatatatat	atatatatga
71761	acacatatat	taagtatacc	atttgattat	ttttgacaaa	tgcatatacc	tggtgtacaa
71821	agtcctatta	agatacagaa	tgtcaccgct	atcccagaaa	gttcccacat	cccactggcc
71881	agtaaatcct	cccctgcgcc	tcccagaggc	agccgttctt	ctgatttttt	tccccatcac
71941	aaattagttt	tgtctcttct	agaacttcat	ataaatggaa	ccatatagca	tacacttgta
72001	ggcctctctc	actgagcata	gtattttgag	atttatccat	gtgttggtg	attcattagt
72061	tgttacctat	ttagtgtctg	gtagtattcc	attgtatgca	gagatcacag	ttgttttacc
72121	atccttctat	tgatagacgc	ctgagctggt	ttgtttgtgg	ccattatgaa	taaaacttca
72181	gcgtacgttc	ttgtgtaagt	ctttttgtgg	ctatatgtat	ttatttctct	tgggggaata
72241	aatagacata	gaattgctat	gtaagttag	ttttacaaga	aaccgccagt	catttttcca
72301	aaatggctct	actatttgta	ctcccaccaa	taatgtatga	acatttggtt	gtaccacatc
72361	ttcaccaaca	tatggtgtag	tcaactcttt	taatttttag	cattctagt	ggcgataaat
72421	ggtatctcgt	ggtttttagt	tgcttttgcc	tgatgactaa	tgatgttgaa	cactttttta
72481	ttatgtgctt	atgctatttg	agtataattc	ctttgtgaag	tatctattaa	aatcttttgc
72541	ccatttttga	ttaggtgggt	gtatatccta	gctgccagtc	ctttgtcagc	tctatatttt
72601	gcaaacatga	aaaccagtc	tgtagtttgg	ctgtttgtta	tgttaatgat	atcgtttagc
72661	caaagttttt	aattttgata	aagtagaatt	tagcagttgt	tttctttcat	ggttattgct

72721	tttctgtatt	gtctctaata	aaccattgca	cgttcccaag	gcacaaagat	attctcctgt
72781	gttttcttct	aattacaggt	ttgagctttc	acttacaggt	ttatgttcca	tcttgaatta
72841	attcttatgt	gtaatatgag	gtgggagatca	aggttccttt	ttcccatat	agacagctag
72901	ttgctttaac	atcacttctt	taaagatttt	ccttccctat	tcggattata	tcacaccttt
72961	gttaaaaaatc	gaaggactca	gtaaatgtgg	gctgggctct	tttctgtttc	atcgatctgt
73021	ttttcaatcc	ttatgccagt	gctacactgt	tttgattact	gtggcttttt	agtgtatctt
73081	gaagtcaagt	aatatgagtc	ttctaacttt	gtaattgttt	ttcaaaattg	ctttagaaat
73141	tctaggtcct	ttgcatttct	atgtaaaatt	tagaatcagc	tggccaatgc	tctattaaaa
73201	agtataatgg	atttagaatt	gtgttaaaac	tatagaacaa	atggaaagaa	ttgacaattt
73261	attgcttctt	gcaattcatg	aacatagtct	atctccttgt	atacttaggt	ttttaaattc
73321	tcttagcaat	cttcattggt	gagattgtat	aagccttttg	taaacaaatt	ctttcaaaat
73381	atttgtatgt	gttttgggtc	tacagtaaat	gaaatgtaaa	tttcattttt	aaattttatt
73441	attattatta	ttattttttt	ttttttgaat	cggagtcttg	ctctgtcgcc	caggctggaa
73501	tgcagtgcg	tgatctcagc	tcactgcaac	ctgtaccctc	tgggctcaag	acattctcct
73561	gcctcagcct	cccagtagc	tgggattaca	ggcatccacc	accatgcctg	gctaattttt
73621	gtatttttag	tagagatggg	gtttcaccat	gttggccagg	ctggctctca	actccggcc
73681	tcagggtgatc	cgcccacttc	ggcctcccaa	agtgtggga	ttacagggtg	gagccaccac
73741	gcccggccat	aaatttcatt	ttttcaaatt	tttgctgcta	atataatac	atacggttga
73801	tttttatata	ttaatgttat	gtcataagac	cttactaaat	tcactactta	attctaaaag
73861	ctatttttgt	aaatccttta	atatttactt	cctaaacaat	catgtcatct	gcaagtacag
73921	tgatttttac	ttttcccttt	tggattgtga	tgcttttctt	tctcttgctt	tactgcactg
73981	cctaggacct	tttcttacag	tgtaaacag	aagtggtaag	agtgggctgc	tctgtcttgt
74041	tccagtgat	acagggaaaa	catttttatt	tcagtattaa	gtccagtggt	gcctgtgggt.
74101	tttttatagt	tacatgtatt	agattgaata	agtttatga	aagggtttat	cattaactca
74161	tttgtctgat	gctttctctg	catctattta	aatggtcata	tgattttcct	cctttatttg
74221	gtaatatgga	tcattttgat	ttttttttaa	cattaaacct	cacatgccta	ggaataaacc
74281	tattatatca	tcattctttac	atattgttgg	attcaacttg	ctaatacttt	gtagaggalt
74341	tttgtgtctg	tgttcatata	gggtggtggt	ctgtaatttt	cttttttata	atttgttgt
74401	caggatttcg	ttgttggtat	tagtgtaacg	caggcttcac	aaaacaagta	aggatgtgtt
74461	gttccctccc	ctgttttctg	aaagtgttca	tgtaacatga	atatgatctc	ttccataaac
74521	gtttgctaga	actcaccagt	gaaactatct	agggtcgtaa	ttttctttat	gggagggttt
74581	tagatcataa	ttcagttcat	ttaatagata	tagagctatt	catattttct	gtttcatctg
74641	tgtccatttt	aaaaagttac	gtttttcaag	gaatttgtct	gtttcattca	ttttgtcaaa
74701	catttttggt	ttatgttgcc	ttattaggct	tttaacatct	gtggaatctt	agtgatcacc
74761	cctgtttcaa	ccctgatact	catcatctgt	gttttctctt	tttttcttgt	ttaccgagt
74821	taggggttta	tcaattttgt	tgctcttttc	aaagaagtag	cttttggttt	tatttccctt
74881	actctgtaga	cttctgcttt	tatttttatt	ctactttctt	tccgtttaat	tgtcttctt
74941	tttctagtga	tttaataagg	tataaaagct	tggccaggcg	caatggctca	cgctgtaat
75001	cccagcactt	tgggaggtga	ggtgggcgga	tcacctgagg	tcaggagttc	gagaccagcc
75061	tggccaacat	ggcaaaaacc	cgtctctact	aaaaatacaa	aagttagcca	tgtgtgggtg
75121	cagcagcctg	taatcccagc	tactggggaa	gctgaggcag	gagaatcgct	tgaacctggg
75181	aggcagaggt	tgcagtgcgc	caagatcacg	ccactgccct	ccaggctgga	taacagagtg
75241	agactccttc	tcaaaaaaag	agaaaaaag	cttggccatc	attttagaca	ttttcctcaa
75301	agcactgctt	tagctgaatc	ccacacattt	tgatatgggt	tattttaatt	attattcaat
75361	tcaaaaatatt	ttttcatacc	ctttatata	atgtatttga	tccatggaat	gtataggaat
75421	gggtgtttta	atttccaaat	ttccagacaa	tgaggttttt	cttgatatct	tattaatttc
75481	taattttatt	tcattttggc	cagagaacct	actctgtata	attttgggtg	tttaaaattt
75541	attgagactt	gttttgtggc	ccagcatatg	tggctctctt	tgggtgaacat	gccatgtgtg
75601	tttgtaaaga	atgtgtgttc	tcagttgtctg	ggtgtcatgt	tctataaata	tcagttaaac
75661	caagatggtt	ggtagtagtg	ttcaggtaaa	ttttgttttt	tattcttttg	tagttctatc
75721	aattgctaag	agattgaaat	ctccaagtat	gattgaggaa	ctctgtacat	ctctctcat
75781	ttatattgat	ttttactgaa	tgtattttgt	aaatctgtta	ttaggtacat	acacatctat
75841	gattgctgtg	ttttcctgat	gtatgagctt	ttcaccattg	tgaattacc	tctttatcat
75901	catgagatgt	ccttccgcct	ccctggctct	gcagtttact	tgggtgttaat	ttagccgtca
75961	tgtgcttact	gtttgcccctg	tgtattatcc	ttttccatac	atttgctttc	caccatgtt
76021	tctttatctt	gaaaatgcat	ttcttttagac	agaagtctac	agtaattggc	tctttttttt
76081	tttatccatt	ttgtctgtgt	gtgcccttta	attggagtg	ttagtctgtt	aacatttgat
76141	gtaattattg	atacatcagt	tttaagctgat	ggtttgattt	atatctgcca	gtttaatcat
76201	ctcctcactt	tggttttcag	tagccaagaa	taatagttgt	aatgaatact	attatggtct
76261	aattctttat	aatgtatttt	ttctatatcc	tttaataggg	aatatcttct	taagagaaag
76321	gtagaggact	ccttatatct	agtaacaatgc	cttaagcata	gaattctggt	acttaaaaaa
76381	tgctaagtga	atgcgggtgga	aagaactgtc	ccttaagaat	caggagacat	agtctccaga
76441	cttattttata	ttattatttt	gcattattac	ctggtttaat	ccacttattt	tgtgtgtgtg
76501	atthagtttc	ctcatctata	aaaatgagga	ggtttaggac	tatattatct	ctaaaattat

76561	acctttctgt	catctatgat	ttgaagcttt	ctaatagaaga	gaattttttat	ctaaaagata
76621	tgttcacaag	ttattcttca	tttagcaacc	actttctgac	aatcattttc	tataatgttt
76681	ttatgtatat	aacccttaaa	tttcaatgtg	gatataatat	taaaagaatg	caatatctgt
76741	gattcttttt	ttctttttcta	tttaataatgt	gttgctgctg	gttcttttagc	tacacgaggg
76801	gactacaaaa	tgatagtttt	tgtagccat	gaaaaaaatc	aaacctcaaa	caaaatgtta
76861	taagctgttt	ttatatatct	taagccctgc	cagctaattgt	gatagggcac	aagggctctt
76921	gattagtctg	taagctgcag	tgccaccagg	ttggtttttc	ataggttagta	ctcatttttta
76981	aatcaaaaatt	ctgtcggttac	ttcattttgt	gttgggcgtg	gttaatttat	agaacctcat
77041	gatataacca	ccaatatcga	ggaagcagac	ttgactacca	tctcaaaaaa	agattggggg
77101	taatagttaa	ttttaaatcc	tcaagcacat	caaacatcca	actcagttaa	gtctagagca
77161	tcactagcag	agcattgggc	agaacttcaa	attttattga	ggatatttca	agatagggtg
77221	atagtttaaga	agaggaattt	gttcatgggtg	ctgctgccaa	ggagttagcac	caataactca
77281	aggtgtttatc	aacaattcaa	acaaatctgg	ctgttcaaag	aagttagtaa	actccactga
77341	cttggttttgg	gaccgtagat	ggagaagaga	ttgtatgttg	tgactaacga	aatgaatat
77401	ttcatctctg	taccattttt	tactgtgcaa	tttggtttga	acaggttgag	tatgaggttg
77461	cagcatgtcc	acacaggga	tgttctgtaa	gccattcagc	aacttgagcc	ttgagttctt
77521	gtgagagttt	aagcaggact	agactcaggc	acatccagtt	gcagtaagga	cagaggtttt
77581	gaaagaggag	ctgccacgaa	ttatttgtaa	tgagaggtgc	ccacctcttt	gataccacag
77641	cttcttgagg	tgaaacaaaa	gatagtttcc	agaagataat	aaagagactt	taaaatcagt
77701	gtgctatctt	ccttccatca	gtgtccccct	gggtggggccc	ttagtgaag	gagacagtaa
77761	taatagatag	tgcttctgtc	aaaaggctgc	tttcttcttt	ccagaaatag	aaacatgctt
77821	cccaagaaat	aatctaaatc	tatttatatt	tctgccactt	ctagcttttt	gttctgtagt
77881	catttctctt	tttttttttt	tttttttttt	tctgtatgtc	ttatcccccc	agctagattg
77941	ttaagcatgt	ctgggacagg	aactattata	ttttacttt	ctgaagaata	cctggcctac
78001	tgcttagcac	cttctaaggt	taactttttg	caagaggaaa	gataagcgg	agatgtatct
78061	ggaggctgct	gttgagtgga	agagagacct	cctatgttcc	cagttatgcc	taattcgtt
78121	attccttctt	caaatttgta	attttttttt	aatcaacgga	gaatttttag	tggtgaagcc
78181	tttttagcctc	tagacatggt	tgagcctgct	ggcttccagg	ggcttccgtg	ccagatggag
78241	agacagacat	aaccacaaat	aacaaggctg	actacttgta	gtgtgggcag	gggtgctggg
78301	aatgggagca	ctgagatctc	actgaggggt	gaggggttct	gggaagatgt	cattagggga
78361	gaaatgtatt	ggagctctgt	tttgaaggct	ggtagccttt	gggttttgg	ggttttttaa
78421	ttcctggggt	attgtcccat	ttctcttcac	ccctgctcca	caatttttaa	aaattctgtc
78481	aacgtaaggt	tttgacttaa	gctatgcttc	acagagaaca	tagcatcttt	tatagctggt
78541	gccactccta	tggcatacag	aaaagcgtag	gattcaacat	aacccccacg	tggtgagctg
78601	tacaggctta	aatgaacct	gtaataccac	aaaaagagca	gtggaattgg	aatcagagga
78661	tcccttttga	tgcttggcct	tgctctcgca	gaagttgaca	gagttgtgct	atgccgaaga
78721	acttgtagcc	cccaaagagg	tatgggaggt	gaactaggta	ctgaaggaga	gccatgcttt
78781	tggtatggctg	tctaagtggc	acatactggt	ggatagctac	ctaaatgata	cagggggatc
78841	atattaatac	caggccataa	agtgtcagca	cagattgggt	gaaagcctgt	atgcgcatat
78901	ttgcatatta	aagaacagtt	atgttgatat	atttacaatc	ctgttaaatg	aagaatcaga
78961	tttggtgaga	tgtaggatta	gatatagtat	caaaaaattt	tcatgagaat	acagtaagcc
79021	tatgggaaaa	attcattggt	ttgtcattca	aatgtgatac	aaatttctgt	ttaattgctt
79081	ttcagatgMa	aatacaattc	tggaaagagg	aaatatagtg	aaagcctttt	tataaatata
79141	aaatattttt	tcaaaacttg	agttgttttt	ttcttccacc	atttctgttg	ctgcaaaagg
79201	taatgacatt	tcccgctga	ggaaaaacat	ttttgaaatg	gagttaaatt	attatttgag
79261	aataaggttt	ccttctgtgg	cctgtactat	attctgatga	cttataaaga	gaccttgtgc
79321	agcagtgtcc	tctgtatgtc	cttggtgggg	cctttgacac	tctgcaatag	gaagactagc
79381	taataaattt	tctctttcct	ctacacaaat	tcccatttgt	gcaattgcct	cccctctagt
79441	ttttcccagt	gatgtaagct	actattatgc	caccacgctt	gctaacttaa	tggtatcact
79501	tttaagaaat	actgcaggtt	ttttaaaaat	accagttgga	Rtgcctcatt	ctaggatatt
79561	ctcattttact	tttaaaaacg	tcattttagag	ccttttgctt	tctgggggtt	tttggtactt
79621	ttttttcaac	ctttgtatgg	tggtctttct	ccataatata	tgaatattta	tttttatttg
79681	aaaaatgttt	tccctcaaac	ccaataattg	atgctggagg	aaggtgtgtt	acgtctctcc
79741	tgtaggcata	tgtagctgta	ctgcgtgcct	tagtaccac	ctgtttaagR	ggcacaggac
79801	ctgatgattc	tattgtttca	gagaagccaa	attagtttgg	atcctgtctt	aggcaagatt
79861	tgataagatc	tgagcccttt	ttctgtctca	tagttcatct	tttagtgact	ttgaactagt
79921	tgttttacctt	tagtctgtgt	gcctgtggta	atcctacctt	aaacctcgcg	gaatacagaa
79981	ataaacaatg	caaaagaggt	caaagtgtcc	gaatgaaaga	tttgagaatt	tggtctctag
80041	attgggtgatt	cttagctttc	tttaattatta	ttggcaataa	ataagtcctg	ttttattgac
80101	aattaagagt	gaaagagtga	acctctgga	attagttaca	aaaattacag	aaattttatg
80161	attttttactt	ttcagtaaa	aacaatgaac	ttcttcagaa	agaaggaaaa	taaattggga
80221	attttttttta	aaggcttctc	tcatcctccc	cccaaaattg	agaacattgt	agaaggggtg
80281	actaagaaga	atgatgggtt	tctaagaatt	gagagatgtt	ggccgggcgc	agtggtctcat
80341	gcccgtaatc	ccagcacttt	ggggggccga	agcaggtgga	tcatttgagg	tcaggagttc

80401	gagaccagcc	tggccaacaa	agtgaacc	cgtctctact	aaaaatacaa	aaaaattagc
80461	tgggtgtggt	ggcacgcacc	tgtaatccta	gctactcggg	aggctgaagc	acaagaaatg
80521	cttgaaccca	ggaggtagag	gttgacgtga	gccaagatcg	cgccactgta	ctacagcctg
80581	ggcaacacag	tgaactccg	aatcaaaaaa	aaaattgaga	gatgttgaaa	agcagagaag
80641	tctgggggtg	gtccacgctg	agttgtctaa	gcagtggtgt	aactggaata	cagagcaagg
80701	actttgaagt	caggtggacc	caacttcaga	tcctgactcc	gattttatta	gctctgtgcc
80761	cttgaacatg	ctgtttactc	cctttcagct	tcagtttcat	cacttgtgaa	atggaggtaa
80821	taggacacac	ttcatagagc	tagtgggagg	attcagtact	agagtgtgtg	gcacagagct
80881	tggcccacaa	cagcctctgt	aaatgtgagc	tcctgtcgcc	tcacttcact	agccctgcta
80941	gcaagagact	tactttctctg	cctctttggc	cagcattaag	gccactaact	gagacagcat
81001	gaatctaagg	tgggtgctcg	tactctgata	catgatttta	gtaaaagaca	aaaatgattg
81061	gtaattttgt	aaatcattgc	taagattaat	agactagggtg	aaaaggcttt	tgtgtatata
81121	gaagacattt	catctaaaat	attcctataa	tcattacata	tgccctcccc	cctttttttt
81181	ctttcattta	aaatatagag	tgggaaataa	ataatgtaga	tggttttcac	ttatgtgaaa
81241	aatgggtgtg	aggaattgaa	aaatagggtt	tcacagttga	aatcactgct	ttcaaggaga
81301	ttatgggttc	agcttactgc	aggagaataa	taatgataat	gatgataata	acattactat
81361	gctagctcag	tcacgtcaat	cttatgagg	gtagatagtg	tattatcctc	attttgcagt
81421	taaaggaacc	gaagcacaga	ggttaatagc	atccccatgt	gcatagttag	gaagtagcag
81481	agccaggatt	tgaactcagg	cagcctggct	ccagagccta	cactccttac	cactatctca
81541	tactaggagg	acagagagaa	ggcagttttt	tgagacagag	taatgtctta	aaacctgcaa
81601	caggttaatt	attttagatt	tgctctacag	gtattctgag	agaagcctca	caaagtgtat
81661	tgggaattact	attgaaacca	gaccagatta	ctgcatgaag	cgacatttaa	gtgacatgtt
81721	gacctatggc	tgacacaagg	tggagattgg	ggtgcagagt	gtttatgaag	atgtggctag
81781	agacaccaac	aggtaagatg	gtggcagggtg	atcttgacac	agtcttcctc	caagttcacc
81841	attttctcta	cattcatacc	cagcctttct	tccttctgac	cactcttagg	gaaagaagta
81901	tgggtattcc	tccttttcag	agttctttct	tctgtctgtg	ttcttaattc	catccccctc
81961	ttccctcatt	ttcagtcctt	ctctagtggg	tcctttccag	cagcctgtaa	acacactcat
82021	ctctctccct	cttcttgtcc	taagcagcct	tgtccatata	gagagcaggg	gagaggactc
82081	ggggttagta	gtttaaagca	ggagagaagg	ccaagaacaa	agaaaagagt	ttgaataaaa
82141	aggacatcag	ggtaatgttt	aagagattca	ttttgtgga	gcaaatacca	actagcaaaa
82201	agtttggtga	tgccatggtt	gtaaactttg	agaaatttgg	gagtgcaaga	aaagaaggaa
82261	gtaagagttg	ttagagcttt	attatgttcc	ctctagaagt	atatctttta	gttgaaaaga
82321	aaatacccg	gaaaggttaa	ctgtataagg	aacttcacag	ttaattgcta	aattagtggg
82381	gtaggggaga	caattgagag	tttattaggc	agctgagatg	attgaccgag	aattcctaacg
82441	gtgggctttg	gaggaaggat	gtgttgagtt	ggatgttaaa	ggccccctta	ttctgtagga
82501	tttggttagg	tgtgattggt	gtgtgcccag	gagaagggaag	caaaggggaa	ctagggaaac
82561	aagcactgat	cagcatctta	gaaacaacca	agaggaatgc	caggccgagg	aggatggaat
82621	tttcccttaa	ggaatagaga	tctgttgaat	attttttagc	atggggaaga	atctcatgga
82681	tatcatgttt	aattaagaaa	gagtaaccga	ataggatgct	cttatgtggg	aaaaggcttc
82741	aaccagggaa	actgtggagt	gctgtagcct	tcaagggctg	tgctagagag	acagttgtgc
82801	agatggagg	gaaggggag	ctttctgaag	aatacgacat	ctgagttggg	acttcattga
82861	agagtctgag	ttgggtcagg	gcagaagagg	aagattgagg	cagaaagaaa	ataatgaact
82921	cctagagatg	ggagtgccat	agcacatttg	aggaatttgg	agtggcttag	tttgattaga
82981	atgtgaaata	ggagagtgat	aagagggtgag	gctgaaaagc	aggcagaggg	gctgattatt
83041	aggggcctga	ggagtacgcc	cagggaggct	ggacatgaat	ctaattggagc	aagccaccag
83101	cacgcataaa	gctacaatgt	gatggagcat	atctgcagct	tagaaaagatc	gcttacagtg
83161	taaggggag	tcagagaaga	gcaagttgga	agggagatca	ggtacgaggc	tgttacagaa
83221	attaagcagg	aaatactgat	gtcctggaaa	gaagtgaatg	gtggtgattg	gtgactgatt
83281	ggaagtgagg	ataagagaga	atgggaagtc	aaggaggctt	ggagagaatg	tgggtttcat
83341	ttggctacac	taagtttcag	atacctgtag	aatagccaag	tgaaggtttg	cttttagagca
83401	gggttaggca	atagacagtt	cagcagtgat	caaaatgttc	tgtgcctgca	actgagtttt
83461	ccattttta	tcttctaaat	ataaatttta	ataaccacat	gtgcctagt	gctgtcattt
83521	taaatagtgc	tgcttttaag	tattggttat	ttgggtttgg	aattcatgag	agaggcctgg
83581	ccttgtccta	ggaattggga	agtcacacgc	aaaaagacat	tagaaattga	agccatggaa
83641	atgggtggga	gcattctgag	agagtctgag	ggagtctgtg	gtttctagga	tccttctgcc
83701	gtgaattggg	acaccttcgt	ggcagaagga	tagccacaag	acttcagttt	agttgtccca
83761	aactattgtt	aggacttggt	tcctgctgga	ttggctctag	aatgccaagt	taagatgatt
83821	tatccacgtg	accccaaagt	actctgcctt	ctcactaagt	cattctctct	ttctactgtt
83881	atttaaggat	cctccaaacc	caccagtgaa	atctattttg	tatcataggt	ttcactcttc
83941	ctcctgtct	acaaataatt	tcacaagatc	tttaaaggaa	agaacatagc	agcttgctgc
84001	ttctcctagg	ttttgagctc	attttaattt	tttaacacag	tattttacatt	cataagctta
84061	ttttgactgg	tgatgctaaa	ccaaataatt	tttaacacac	tttaactttt	ccatcccaat
84121	gctttacttt	caagacttta	agtagatgtg	taagagaatt	tctgagaaat	atctcagaaa
84181	taaagttatt	tacttccagg	cctcctgaga	ggtggggatg	taaaaaatga	ctttctgaga

84241	tgtctttttcc	cccagccata	tgatttttgg	tatggaactt	ctgttacctg	tctcagagat
84301	tgttttcccc	ttgttccttg	tgtaactttt	gtcaccatt	tatttgtgta	acagatgtta
84361	cacaaatagg	tagtgatttt	cacaggtagt	gattttcttt	ccttaggggc	cacactgtga
84421	aggcagtgtg	tgagtcat	cacctggcca	aagattccgg	ttttaaaagt	gtggcccata
84481	tgatgcctga	cctgccaac	gtgggactag	aaagagacat	tgaacagttc	acagtaagt
84541	tgacttcagc	caggcgcat	cagaatggct	ctgcatgttt	cttatcccat	ctggtcttgt
84601	tgcttgttca	ctgttgatgt	tttccagtg	ttaaagaaat	catccttatt	atagaatatt
84661	agaaacacag	tagggtaaga	gacatcacc	atagcccat	cagacaaaac	tcttaacatt
84721	ttaatgaatt	ctttccaaga	ttttctatg	cataggtttt	ttggggggtg	aggttgttta
84781	tgcaaatata	acacagacat	acaaatata	atgacatagg	aaattattgt	aatagttttt
84841	tcacttaaca	ttttaacaag	catttatcca	tggtgtagcc	tggtctctgt	taacatgtta
84901	acatatattaa	atgtcttttt	aatttaatat	gtcatacctt	gcttaaacc	ttgcaagttt
84961	ttcagaat	aagatatact	ggttttttaa	tattatttaa	aaaacactgc	aggtaaata
85021	ctttatgtgt	ataat	ttagtattta	ggtttatctg	tttatgataa	attccatgga
85081	gtagaattat	caggtaaagg	tagaaacatt	ttaagaatct	cactgtatta	ataagttgct
85141	ttgcagagag	ttataacaat	tttactcat	ttcagcagag	ttttcttcc	caaataattc
85201	tctcctgttt	attttctata	tatgcttata	gatgcata	cacatatgta	catatat
85261	actgtttgct	agtttgatca	ataaaagtgc	cattttaact	tgcattttga	aattgtaata
85321	atacat	caagtttgta	tttactaatt	acatttctt	catctttttg	gaattgtgtg
85381	tttttatcct	ctgactcatt	attctattga	gataagaata	tttttagtat	tagtttatat
85441	gaatttttta	ggttaaggaga	ttatggactc	attggttttt	atatttgttt	taaata
85501	ccttaaccta	ttgttgacct	tttactgtg	cttccattaa	gtcgatcttt	ttttccttt
85561	tttttttttt	tcttcaaagt	cctctaagct	tagaactgcc	tttctgttct	ttgaatttta
85621	actctgggtt	gttttattgt	tgttatttta	actctcacag	tttcttgatc	catgttaact
85681	ttgttgtggg	aaatgagatt	ggtcaggcct	tttgtttttc	ttcagattct	ccgatagggc
85741	aagtgtgttt	taagatgatg	acgatattta	tcaagttatt	tcactcataa	aagtattttt
85801	aatttgcctc	gttttagaaaa	atgcagttta	ccatgtctta	tttgtgaagt	taataattct
85861	ttcctctcaa	gaagtttact	tttattattt	tttaagctgg	tgaactgcac	ctgaggacct
85921	gtttatgaga	taacattaat	tcacagatat	ccctctctgt	aattccccca	caatggatat
85981	gtaatagact	gtgagtctag	tatcctcccc	tcagtaacta	atagccaggt	gctcagacac
86041	caggccaaaa	aggcacctga	aacacacttt	ttattatcat	gattattatt	attattatta
86101	tactttaagt	tctagggtat	gtgtgcacaa	cgtgcaggtt	tgttacatag	gtatacatgt
86161	gccacgttgg	ttggctgcac	ccatgaactt	gtcatttaca	ttaggtattt	ctcctaattgc
86221	tgteccctcg	cctgcccccc	acccacgcac	aggccccggg	gtatgatgtt	ccccaccctg
86281	tgteccagtg	tgaacacac	tggtcttttag	cacagctcct	aggagagacca	gtaggatagc
86341	cagaactcat	aaagtttaag	tgctcctgaa	aaatatagta	tgtttcatac	tgtaggaagc
86401	catagcaaat	agctgagctc	cagtacatat	tttcttgggt	gtcttcaaca	taagggaatag
86461	ttcagtaata	acattgcaag	gacaaccttt	cttgtaaaaca	gatttatttg	cttgtttgta
86521	gtctttgcag	ggaaactcac	agggaatatg	agtggatcat	atttcttttc	taactgcac
86581	ctagtaaaag	gtaacagaag	ggagctaaaa	tgaanaatct	gagttctttc	tagagaagaa
86641	ttctctacaa	attaaattgt	tttaaaataa	ataagtttta	gaaattgata	agaggcaaac
86701	agctcagtag	gaaaatgggc	aaaaaacttg	aacaagcatt	tcacaaaaga	gaatattcaa
86761	atagccaata	aacctatcaa	aaagtctcca	aagtaattaa	tcttcaaaaa	agcgaaattt
86821	aaaactacat	ttgggtctcag	ttttctacag	cacctgggtc	ttgtattgtg	gatataattt
86881	atcttttctt	atctctctga	agaaaattat	tacagatttt	ttttatacat	tttatttttt
86941	tgtatgcatt	gctctgtttt	cctccaagtg	ccttttggtt	tctttttgtt	ttggttgcc
87001	ccttaatgta	cttggttttc	cccaagttta	tgatgatctt	caaccagttt	cttatgcctt
87061	tttttttttt	tccttgagac	ggagtctcgc	actgtcattc	aggctggagt	gcaatggcac
87121	gatcttggct	cactgcaacc	tctgcctccc	gggttcaagc	aattctcctg	cctcagcctt
87181	ctgagtagct	gggattacag	gtgcccacca	ccacacctgg	ctaatttttt	gtatttttag
87241	tagagacggg	gtttcaccat	gttgaccagg	ctggtctcga	acttctgacc	tcgtgatccg
87301	ccgcctctg	cctcccaaag	tgctgggatt	acaggcgtga	gccactgcgc	ctggcctaaa
87361	gcgattcctt	atgcttttaa	gagttaggggt	ttgaaaagcc	ggatggcaag	gcctgtgtgc
87421	ctagcttgca	ggcgtgcttc	actgaagaat	gttcttttag	caataagcgc	ttctttcttt
87481	atgggattcc	taaagtctcag	tatcgtgaac	tcttacgtag	agccatttga	ttcatccaga
87541	gatgaactct	ccaacttctc	gctttgggct	gagtggcctg	agtatatctc	aggaagtgg
87601	ttgctgacat	ccagagagta	ggcgagagag	tccaccagtc	tatgtgctgg	cttttccctg
87661	ttttcagcct	tggtcttcat	ctagaccttc	ctttgtacct	ggaatctcag	agcgagagg
87721	ctttccagtt	tctccaggga	ctaaatgttc	tcactcgcct	gttcgggaaa	ggctaagggt
87781	taacggataa	ttccctctat	gtacacagta	gatccagtac	ccctgttttt	atttccatgt
87841	ctctctccca	tctgcccag	ttcttggtac	ctctgatccg	tgagcctttc	tggagactctg
87901	cagtgtgatt	gagcaggcct	ttcgttactg	tctccttttg	ctattttctt	tacttgcaaa
87961	gtcattcaca	agccttccta	gtttctgtct	ttgaagaatt	tggtgccatc	ttgggtctca
88021	ttcttgccct	ctttcttgtt	ctctttgttc	ttctgggttt	atacttttta	aaaaattctt

88081	tcctgttcct	tttttttttt	tttttcttga	gatggagtct	cgctctgtca	cccaggctgg
88141	agtgcagtgg	cacgatcttg	gctcactgca	acctccgcct	cctgggttca	agcaattctc
88201	gtgcctcagc	ctcctcaata	gctgggatta	cagacgcgca	ccaccacgtc	cagctaattt
88261	ttgtattttt	agtagagatg	gggtttcacc	atgttagcca	ggctgggtct	gaactcctga
88321	cctcaggtga	tccacctgcc	ttggcctccc	aacgtgctgg	gattacaggt	gtgagccact
88381	gcgcctggcc	cctgttactt	tagtgtggtt	ttggagggaa	agaaaataaa	cactttttta
88441	aattctctag	tgtaactgaa	aattgagaac	caatcaattt	tcattctttc	caaaggaatg
88501	tatcagcctg	tatgtctccg	aaacttatta	tacaagcacc	agaaataact	ttctaaatcc
88561	catgttcaaa	tgtagattca	gactgggtga	acaatggctt	ccatacgtgc	tctttgtggg
88621	atggctgcag	aataatttatt	aaagacctga	tgattcttta	aaatataaat	gttacaggaa
88681	aatcacacaa	ttacagaaga	agatctctga	aactaaatag	tacataaaga	atacagcatt
88741	taagaacgtg	acagatgtca	atgagagatg	gaaaagtcta	ttgttatatt	tctttgtaaa
88801	gcaagttaaa	caatgaagag	ctttggaaga	ctaattgtgg	aaagataaaa	agtatttttc
88861	cagttttacaa	gatagagccc	tgccagctga	gtagatactc	ccaccaaggt	gggagtgaga
88921	ggggactgtg	agtcacactt	catttctcca	aaatatatta	tgcaaaaaca	gattagagca
88981	aattatcagt	gcaaagtcac	ggaaaggaga	atgcctgggtg	tctcattgct	acatcccaaa
89041	ataaagaaag	aaagcagcca	gctggggtga	tgtagtataa	gaaatgactt	gccaaaagta
89101	gttgagttag	atctattgct	cagtgtccag	ataacaaatg	gacaagggtg	agcccaggat
89161	aggaatgggtg	ccagttgggt	tagggacaga	gtagtcatat	ccagggatca	gcaaactttt
89221	cctctaaagg	gccctatagt	aaacatttta	aattttgtaa	gctagatgat	ctctgtcaca
89281	acttttgaaa	tgtaagcaat	taagtatat	ccattattaa	agaaagggtat	actclagagg
89341	tgaaggcaga	cattgaatag	attgtcatgt	gatgataaga	aagaggtaat	actgggttca
89401	gtgggatttt	ttttaagtgg	gacagctagt	atttgaaagt	cagagggggc	tctctgaaga
89461	agtggcattc	aaactgaaac	ctgaagatta	gctagataaa	gaaaaattga	tgaactttcc
89521	aggcaaagga	aattgccttt	gcaggagtgg	aaaggccaga	tggtgagggg	tggcagtaga
89581	tgagattgtg	caggagacaa	gctgggaatgg	tgagggccta	gtgcagtcca	gcacacactt
89641	gctttgcccc	agtgagacta	cagaaacaag	gagtttctgt	tctgtctgta	cgccctacct
89701	ggtcagaagc	aaaggctgcc	ccagggccta	ctgggtgtgc	cagagaagct	gtcaggggtt
89761	gagatttcac	cctcggtgat	ctctgcataa	ctaattggaga	agtcattttc	tgttctctat
89821	tcacaggagt	tttttgagaa	ccctgctttt	cgtcccgatg	ggctgaaact	ctatcctacc
89881	ctggtgattc	gtgggaccgg	gcttttagag	ctttggaaat	caggaagata	taagagttac
89941	tctcctagt	acctgggtga	attgggtggc	cggatcctag	ccctcgtgcc	tccatggact
90001	cgagtgtacc	gagtacagag	gtagtgtggt	atcttttatt	cctaaaaatag	ttggtgacta
90061	gtctgtttac	tatttctcat	ggaatatgtc	tgatttcata	ttgaggggtt	tggatttttc
90121	ttaatggaaa	taagataact	ggaatgctat	ctgtaaatag	ggagggatgg	aaatcatagc
90181	atgtctaagc	cactttgcc	ataacgtatt	tatttatcta	cccattcatt	catgagcctg
90241	gagacagagc	catgacggtc	aataggcatg	gtgcttgctt	ccgagcagct	tatggtctag
90301	ttcagtggtt	cctcttccag	gtctgcttcc	atctagatgc	agtaatgggt	atgagcataa
90361	gaagtgtggc	cgtgtgtgca	atctctgttc	tagagcctct	gaaagaaaaa	gtagcaacaa
90421	tcactcttta	cagatatata	tgtaaatgtg	gaggaaagg	gacatatatt	tgatggcttg
90481	aagaaaaaca	aataatctga	actgctttct	tcctagaaaa	gagaaagtaa	gatctcattt
90541	acaatcagga	accttatcta	cctattttata	cttacatata	tacatacata	tatactggaa
90601	acataacatc	acagaaaaat	tttagaagcc	ataaagttag	ccgtatacct	acaatcctaa
90661	caaagccaat	tcccaacac	acattcccac	caccttgcca	aaaactaccc	aggttccaat
90721	attacaaaatg	caaggccaga	aggagctgc	aaatgcatta	caatcagctg	ctagagcagg
90781	actccactga	gcacagaaac	tgtgtgactg	atcatgtaaa	gcaatgaaca	ctgaaaacaa
90841	gcYgaattaa	ctacttaagg	agaattatga	aggatataaa	gtaactgact	tgcttgaagg
90901	tcaaaggaca	tttacatgat	acttctgctg	catactgatt	tctcagtttt	aaaatcattg
90961	gccaaaactgc	aggatctRaa	ttgcctatat	ggtctctatt	tttaaaaaata	cacctaaaga
91021	tactaaggag	attttaatatg	aaaaatcaac	tSttgaaatt	gcttggtgtg	cccttcccct
91081	tggtgttcaa	cccttctgaa	gcaattcagt	ccaaaggaca	ttaggtgggtg	YRgagcagt
91141	taggtatcca	catgcaggag	gcagggagcc	acagggtcca	gagcagggtg	agaagggtcac
91201	tcattcacat	gcaggaggca	gaggcgtggt	gcaggctgat	gaagtacaga	tgtggcgaaa
91261	agggcattca	ttcccaccaa	gggcagatct	ggtacaggat	gtcagagctg	tacgtcctct
91321	aggggtgggt	tggggggtata	tgcaagaagag	gagacagcaa	gaaaagacca	gttacttaca
91381	gggagttgat	ctaactagca	gatatattaa	ggatactggg	tgctagggtt	tttttgttat
91441	cttagaagtc	aattacaaaa	gttgaaaaag	gagaaaaatta	gtgtgagcac	tgtggtgttt
91501	tggaacagga	gatagtgggtg	tgaactcatg	gtttccaaca	tataggtaca	tgtagaaata
91561	agtgtaaatg	taatgaataa	caacacaggt	ggcccttcat	atcttcggat	tctgcattct
91621	acaaattcaa	ccaactgcag	atggaaaaata	ttcagaaaaa	acagtgggtg	gttattgcat
91681	ctgtactgaa	catgtataga	tttttttttc	ttgtcattac	tccctaaaca	atacagtata
91741	acaactagtt	atagtactta	cagtgtatta	gatatcaata	aataatctag	aggccaggcg
91801	tggtggctca	cgctgcaat	ccagcatttt	tggtgggccc	aagtgggccc	attacctgag
91861	gtcaggagtt	cgagaccagc	ctggccaaca	tggtgaaacc	ctgtctctac	taaaaatata

```

91921   aaaaatggcc agacgtggtg gcaggtgcct gtaatcccag ctactcagga ggctgaggca
91981   ggagaatcac ttgaacctgg gagcgaggag ttgcagtggag cagagatcat gccattgcat
92041   tccaacctgg gtgacaaaag tgaaactcca tctcaaaaaa aagtaatcta gaggtgattt
92101   aaagtataca ggaggatgtg cataggttat atgcaaatatc tacaccattt tatatgaggg
92161   acttgagcat ccatggattt tgggtgtctt gggagtccta gaaccaatcc cctttggata
92221   ccgagggaca actgtacata ttttgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt
92281   gtgttttata gcaatcagtg tacttagttt ccagtggagc cagctggttt ctagatcctg
92341   gtctctaaat gctctccatt actaaaagga acctcttgtg ccagaaagta aggaagtatt
92401   caaagaagaa aagggaatct ggcttaaagg gctcccattg gccaaatctg ggacaatttg
92461   agcatcaaaa taaatattga tattaacaga ttataaccag ctgaataaaa taggaatcca
92521   ctagtttata ccaatattcc taattaataa attgtaagtg tgattcagaa acataatatt
92581   tgcagtcagt cttcatcatt ctacagattcc ttatttgcaa attttcctac tcactaaaat
92641   tcgtttgtga ctcaaattcca atactggcat tttcacagtc attcttgggc atgcttagca
92701   gtgaaaaatt tgagttgccc agcattgtca ccagctgagg tcaaacaggt ggtgctctgc
92761   cttctcattt cagcgctcac actgggagcc aatgtccttc ttgcagctctg tttcatgcca
92821   tgtttttcaa agttttgtgc agttttttga ggattcgcca tttagtattg ccccaagca
92881   tagtactgaa gtgctggcta ttgctcctga gtgcaggaa gctgtgacgt gccttctgga
92941   gagaatacat gttatgtaag ctttgttcag gcatgagtga tgtgccattg gccatgagtt
93001   cagtgttcat gaatcaacaa tatatatcag ataaggtgtg tttaaacaga aacacacata
93061   aaacaagggt atgtattaaa cagtcgacaa aaatattgtg accagggcct caggggaacc
93121   taaccttgta tttcccttaa gagcaattac tcagtataca gtgtaagtgt gttggtgaca
93181   ctcggtgtag aatgttaact acttattgcc tgctaacaga actagaacca actaccgact
93241   gtacatagtt gcagagtatc tccccacca tacttagcaa ttgcagagtg aaagagaata
93301   acctaatagt gggaaaacct ggcagctgat caaagtgaac attatcagta ataggacaaa
93361   tcacaattct aggtgcctg atagagggtg tgagaacgta gcatacttc tctgttactc
93421   agccagaact tcacgatctg aatccaatag tcaggagacg ttagacaaaa caaaattgag
93481   aaacatttcta aaacataact ggcccgtaat cttcaaaagt atcaggatga agaaaagagt
93541   tctaacttga aggagactac acagacatga caattaaata caatgttctg aactgggtcc
93601   tttggctgta aaggacttta atgggacact ggtgaaactc aaatggggtc taggattaga
93661   cactagtga gcatccatgt taacttcctc attttgatgg ttcttttgtg gttatgtaag
93721   agaatacctt cgtaggaaat atttaaagta ttctcagggc gttgggacat cttatttgca
93781   aattaacttt caaatagtctc aggaaaataa ttatttgtac tgcacttggg actttttatg
93841   taactttgtg attgtttcaa aatctaaact aatacactag acaacaaagc tgcagtaaat
93901   agcattcaca gagacaaaag aaaaagacga tttgcatgta taactgactc aagcaaccag
93961   aaagcagtaa agatgcagct tctcagccgc agagcatggc aacagatagc acaaaatagt
94021   taacatttcc acatccagaa ataccgaaag taacaaggca tggatccac aggcaatcag
94081   atgaagaatt tgcattttct cccttcgttt tcggcttaaa tttctttctg ctgtaaccag
94141   tgagtcaaaa gcagcagaga cattgctata ctattaaatg cttaatttta tagactttta
94201   aataccacta tgcgacatta tcttggccta ccaacaactg agaaaataag gacttgaaga
94261   aaacaagggc agcagaaata gtctctatag gaaaaatgta aaggagtgat gacaggaaga
94321   gtggagaaca ctttatgat acaaccgaaa ggtcatatga taagaaaagg ctcactattat
94381   taataagcaa catatgaggg ctgttaagga aacctggctg ttctcagttc ttctgataYg
94441   acccgtttac acttatccta tggagaattg gcccatagga gagtccacaa gctctaattt
94501   caggaggccc aactgtcaca tatgtttctc tatcaataca tcacttatgg agcttgtgag
94561   tgaggagtgg tacaggtgga tgtaggctta agatttcgaa gtgacactct tacttcacac
94621   Wtaggtgggt ggattaaaaa gtaatggaat actgtaaaac aagaggcagt ggtgcaatct
94681   ctttcaaaat gtggaattag taatgtcctc gatgagtctg aggacagttt attttttgaa
94741   gatagtgatg actcagttcc taacacagta aggccattgg tgaagagagt catccagagt
94801   ctagctgtga tgtcagggtt ttatgtctta ggattgataa tgtagttaac

```

CHDC1 genomic sequence (SEQ ID NO: 3)

>13:44917401-45013900

```

1       tcaaatcttc tcaaattttc cccatcttaa aaactaaacc agttttctct gctttatttg
61      ttttttagtc aactattaat cttttcttt cttctagccc ccaaacgctt ggagtaagta
121     acttacacct actctttcta ctttctttg cttctcctta cctctgcagt ttggatgttg
181     ataaactgtc ctagaactac tctctcaaag acagaggatt ttctgtgtcc cccaagtctt
241     tgYctcttgc atttgctatt attgcaggac aaggacagag ctctctctgt tgcaactcca
301     gtaccaaaaca ttacctgtga gttagaagcg gttcagcaaa tctttgttga atacagactc
361     tgtgggtctt tcttttcaaa aacttcttgt ccatccatcc ttatgcatgc cctgtgctcc
421     atcctagtcc aagcaccctt tgtcacacct cctataacct gtaacactat tctggcttgc

```

481	ttcttcaatt	ccaattcatt	ctgcagagta	tcattttcatt	tacatcttcc	ttgggactca
541	cccatgactt	caatgactat	tgctgcacaa	tcaatcttct	gaattggtat	ttttaaaatg
601	taattaatta	actgatttat	tattatattg	attaattata	gatgcacaag	gtaaaactga
661	acaacacagat	atctagagaa	agagaagtct	ttctctcatc	ccattctacc	agttttcttt
721	cccataagcc	atcactgata	gcagtttctt	gtccctcctt	gcagagactt	tctgtgctag
781	acaattgctt	cctgattgga	ctatggctca	ccatcatgcc	tggggccctg	tcttgaatcc
841	tagccagttg	ggctagggtc	ttgtttacct	cctgggcatt	ctcccatgtt	gggctgtgcc
901	atacatcccc	ccaggcaatg	atttatgatt	tgcttcaatt	ggcatagtct	aagatgatcc
961	ttgctcttag	aagagtatct	gggactccag	cggctgtgac	cagacatgca	ggtaacactc
1021	ccctgtgggt	attatgaggt	attccaacac	tgttcaccct	cactgggggtg	gacttccttt
1081	ctggtcagtc	ctctatccca	accacgctaa	ttctggcttt	actctctctg	cagcaaccag
1141	aaatcgaaga	aagtcctaac	tggttttact	tttgaactgt	ggtgagttgt	gctacgtgtt
1201	gacctgattt	gccaacagat	aaaatgagat	gacatgttct	ctttgcttct	gcactattgc
1261	ttcactgtct	agcaggacct	ctctttagat	gtcgcactgg	caggatttga	gaaataactt
1321	ggggcggttag	acagctaattg	tttaaaccct	gtgtatgtga	atagtgatag	catctgggta
1381	tgggaagctt	aacgccattg	tagccagagg	gcattttctg	ctaaatctgg	agagaacaac
1441	cactctggga	gcctctgact	agacccta	gacagcctta	agttcggaaa	atatccctgt
1501	aatgatata	agagtcttaa	aaacggcctc	tggggaacat	gaaaattact	gctttagttt
1561	gatgcaaagt	accttaacta	caaataaggaa	gacttttttg	tcctttttaa	gtccaccta
1621	tgaacaacag	gcctgaaaag	gtacttataa	tattgatttg	gaggtaacat	ttgatcagtt
1681	gaattcgcat	ataattggcg	atcgaaagat	gacacatcca	tgagagagac	taccccagtc
1741	cagggtttct	caaagtgtga	gccatggata	acagtcttgc	aagatgagat	ttattggtca
1801	aatatatagg	aaaaaactgc	agtgaatcat	ctttccttct	tccatgttca	caatgtacat
1861	tagcatatca	aggattcaga	gaagtgtgtt	agtaaaacaa	cttctaaact	tttaaattaa
1921	ccctttgtct	cccaaactta	tttggcgaaa	aaatgtttta	tgcatagtgc	ctggttagcag
1981	cctaagaac	taaggttcag	agacacacac	actggataat	gctgcttcag	tttattcatg
2041	caagttaaat	gatttacggt	catgcatgac	acaatgacgt	tttggtcagt	gacagactgt
2101	atggaatgat	cgtcccataa	gattataata	ccaaattttt	actgtacttt	ttctatattt
2161	agatatgttt	agatacccaa	atgcataact	accacggtgt	tacaattgcc	tatagtattc
2221	agtattgtaa	catgctgtac	aggctagctg	caataggcta	taccatata	agcccagggtg
2281	cataatagga	tatacatctt	agttttgtgt	aagtatactc	catgatgttc	acacaacaag
2341	aaagtcattt	gaaatgtact	tgtcagaata	tatccccatt	gttaagagag	acatgactat
2401	attaataaac	aaataactgg	ttcctctaaa	taattcacat	attcatcata	ttacaaacca
2461	ataatacaat	atgcctatgt	tataaaccaa	taacataata	taatgaactg	gtgaaacatt
2521	taggtgaacc	agttgcaata	aagctgtttt	tttaatttat	ctggttatcta	acatttatcat
2581	ttgctttttt	tttaaagtca	ggtatttaaa	aatcgatttt	cccttttttt	tgagatggag
2641	ttctgtcttg	tcaccaggct	ggagtgcagt	ggcacgatct	cggctcactg	caatttccgc
2701	ctcccgggtt	caagcgattc	ccctgcctca	gcctcccgag	tagctgggac	tacaggcaca
2761	tgccaccatg	cctggcta	tttttttatt	ttagttagaga	cggggtttca	ccatgttggc
2821	caggatgggt	tcgatctcct	gaccttgtga	atccgcctgc	cttggcctcc	caaagtgtcg
2881	ggattacagg	gtgagccac	cgcacccgcc	ctaaaaatct	attttcttat	aagttaaagt
2941	cttattagtt	gatgcctttt	ggtccttatt	ttctagtga	ttcacctgga	agcctcctaa
3001	acctctggca	ggaaccagag	ggtttgtctac	tgctcactga	tgttttttcc	ttacaactta
3061	cttaaggaga	tctcaacctc	cttaagccaa	aatattcttt	tattattatt	atttttattt
3121	caatagggtt	ttgggggaaca	ggtggtgttt	ggttacctga	ataagttctt	tagtgggtgt
3181	ctctgagatt	ctggtgcacc	catcacctga	gcagagcaca	ctggacctaa	tgtgcagttc
3241	tttctctctc	accacttccc	actctttccc	ccaagccccc	aaagtccatg	tatcattcta
3301	ttttttcttt	tttttagacag	agtcctcactc	cgtcgcccag	gctggagtgc	agtgggtgca
3361	tctcagctca	ctgcaacctc	cgtttcccag	ttcaagtgat	tctcatgcct	cagcctcccc
3421	agaagctggg	actccaggcg	cacgcatcga	caccagctca	atttttgtat	tttttagtaga
3481	gatgggtctt	caccatgttg	gccagactgg	tctcaaaactc	ctgacctcaa	gtgatccgcc
3541	cacttcggcc	tcccaatgtg	ctgggattat	gggctgtgaa	caccatgacc	agctgttcta
3601	tcattcttat	ccttttgcatt	cctcatagct	tagctcccac	ttataagtga	gaacatatga
3661	ttgggttttcc	attttttagt	tacaaaata	ttcttttcat	tacaatcagc	ttctaatacca
3721	cagtggtcca	aaccaagt	tctcaggtt	ttttcaaaat	cctattatca	ctcataagtt
3781	ctaggtcaga	atgtgtctc	tcogttatac	aattgtttat	attatctcat	ataatttctt
3841	aaaaagcatt	gatttaaaaa	caaaaatcac	acataaatga	aaattgacat	ttcataagac
3901	catgtacaca	tatttatgaa	gacagaacaa	ggactaaagt	ataggttatt	gatttatcaa
3961	gactgtcctt	ctattgtcat	tcagagaatg	gtcaagaagc	cagcttctct	gagaatccta
4021	tcttgaggca	tagattctgc	agtttcttta	ttccgcagga	actgctgtag	aaaggagtac
4081	ttaaataaag	tctccctggg	ctaggagcca	gaactccacc	ttaatctggg	agaaaacagg
4141	gaaataaagt	ggccagaggc	tagtagctac	tgctttgttg	cttcaagaga	aggaagtatg
4201	agcctctatg	aaggggttga	ggttttttgc	tggggacttt	ggggcacaga	atgtaaaaaa
4261	aaaactgcgt	ctggacaata	gtcagctatg	aagtatttaa	aaatgtaaag	gggctgggtg

4321	cagtggtgtg	tgctgtgaat	cccagcactt	tgggaggccg	aggctggtgg	attacctgag
4381	gtcaggaggt	taagaccagc	ctgatcaata	tgggtgaaacc	ccgtctctac	taaaatacaa
4441	aatttagccc	agcatggtgg	tgtgcaactg	tagtcccagc	taattggggag	gctgagacag
4501	gagaattgct	tgaacctggg	aggcggagggt	tgcatgagc	cgggattgaa	ccactgcact
4561	ccagccaggg	cgacagagcg	agacactgtc	ccataaataa	ataaataaat	aaaggggcca
4621	ggcacagtgg	ctcatacctg	taatcccagc	actttgggga	ggccaagggtg	agtggatcac
4681	ctgagggtcag	gagtttgaaa	ccagcctggc	taatgtgttg	aacctgcac	tctactaaaa
4741	ataccaaaaa	aaaaaaaaaa	aaaaaaaaaat	tagccaggca	tgggtggcagg	tgctgtagt
4801	cctaactact	cgggagggtg	aggcaggaga	atcccttgaa	catgggaggt	ggagggtaca
4861	gtgagactcc	atctcaaaaa	aaaaaaaaaa	aaaaaaaaaag	aaaagaagat	tagttgatgt
4921	tacattcagg	taagaaaatg	cttgaaagca	ggtgactaga	cttatgtatg	actaacttcc
4981	ttttaagaaa	accaagggtc	tcagcaagta	cagtagttcc	cagagttttt	gatttcagag
5041	ataaacaaca	tcaaaataat	tggtaactga	ctttagtaac	ctccttttgt	tattttttga
5101	aaaaggatac	ttaaaaaaca	ctacaattta	ttgttaccat	tgttttggtt	tttttttttc
5161	atcagaaaaa	aaaatagtga	aggcacagcc	tcacaattat	atataatttt	tgaattataa
5221	ttcatctcag	ttcatgaaaa	actcaggaca	ttgtcttatt	tttttttctt	cattttatca
5281	gggtcatttt	ttcatcaact	agctgattca	catagaaaaa	tcctgaaaca	catattattt
5341	atcttcagag	aaattcatgt	tgtgttact	agttgcatca	aatatctcag	ccagacaaca
5401	ttatttggtg	aatcttccag	gcagcggatt	tttttaaaag	ggcctgattc	ctgcccttaa
5461	tagcaaaagg	gtgctgtttt	cagaataagt	aaacagggtg	ttaaacaaaa	gagagatgtc
5521	catcttcttg	agaaaatagt	gaagacgcaa	gacagcaagg	ggaaatggct	tgtcaggttt
5581	ctttaatat	tagaaataat	caagatcatt	ttcaagggtc	aagtactctc	tcccatttta
5641	agaaaaactc	cttcattcag	ggaaatcatt	tgaacttta	aagatgtatc	aacaagtacc
5701	ccataaattt	atacaataa	aaaaggacat	atcaacagggt	ggcaaacctg	ttaaacacat
5761	ggtctattat	gcaatactta	attgattact	aatattaatt	gtgactcat	tatatcaaaa
5821	gtactccctt	cgattctgag	gcattggggag	agagttgcaa	agatgaatag	gtccctattt
5881	catggaattt	acagtttagt	aatctgtaga	tgggtaaaca	aaccacaatg	agtaggtcaa
5941	tctaacagag	ggactgatca	tttgttaaaa	agcaaatcgt	gggctcatat	gcaatccatt
6001	attttcttta	aagaagaatt	agtgtctttt	taaaaaata	tgtatgcttc	ctatgatgca
6061	tcacaacgac	tcttagtctg	aactaaattt	gcagtcagcc	agaagatgtc	ccaaatcatg
6121	tcctgactgc	tgtatcacaca	gactaacacc	tggtttcagc	tgaagattct	tttaattgga
6181	tttttttctg	tggtttctgtt	tttactgga	ttttgttttt	ctctgttctt	ctccctggag
6241	gaaagtggaa	atttggttac	ttttttgtga	tggaaagtata	ctttcattta	ttattattgt
6301	cgttattatt	atgaattttg	aagccggact	tttaaaagcc	tggaaaggctg	cggcgaatgg
6361	catgtgagag	gggagggttg	cgggaggata	ccagcagggtg	gcgcgtgtct	cctccgggca
6421	gattgcaatg	ctttgtcag	ggccaagaat	cccgcaggaa	ttaaggggcaa	ctatctgtgc
6481	agtcaattca	agtgcacacg	atcatgccat	acttaagtgc	cgagtccatg	aagtcaccag
6541	ttccagcctg	ttgtctaggc	catcaaaact	atgtattacc	acttaaaaaa	ccagctcccc
6601	tactccaagg	gttgcatcac	tttctgtctt	gacaccccca	caaccccgca	cccacccccc
6661	gactcattaa	ccccaagcca	agtcgtcat	aggggacaat	caagccttaa	gaatagagca
6721	aaaataaagg	ccctggaaag	ttggacatga	agctttaact	tcaggaaagg	aatggatttc
6781	tattaatcgt	ccagctgggtc	tgcccttagg	gtgtcagatg	ttctctgtcg	cccctgggtg
6841	gccagcattg	gcctgggtgc	cactgggtgg	ctgatggcag	aatgtgaatt	cacttgcaag
6901	cggacatgtt	cgtttccatg	ccccaaagcg	ggaagaagaa	gggcacaagg	agacttcgat
6961	ggttctttca	gctcagactc	gagctgtgac	tgatttgctg	tttgggttca	ctggctttct
7021	agcctctcct	ggcaccacc	aatttcgggg	agcagccagc	tcactcttct	cctacgcaa
7081	tgtcttctgt	cccccataaa	cactgccctc	attttgagtt	tttctttcct	atgtctgtag
7141	ctttcttctg	ccgctcagca	aacactctgt	cccactcgta	ttcgtccctt	ggcgtggcg
7201	tctggcccag	gggtacactg	tgaaggaggg	caaccctaga	ttctgagggc	cctagctcga
7261	atctagcgcc	cctctaactg	cagggtgac	cttagatata	tttattaact	tcctctgtag
7321	actcagggga	acaacttttg	cactcctggg	gtccctggat	taatgaggta	acttacgtag
7381	actccctaaa	caagcgtctg	gttcattgta	agcgcccaa	tgccaggggg	catctgcagg
7441	gatgacccca	tagtgaatga	aaactgcctc	ttggggatct	ctgggttaaa	cggccagaat
7501	aaatcgcaga	attgagagaa	ccctctcctg	ccccaaacct	cactgtactg	taaaatcccat
7561	tcattgcctc	ccgcggtttt	atccgtgttc	tcagaccacc	cccctatccc	cgcacccctc
7621	agtcacctaa	agcctcctgg	gcggcgaaag	gacgcccaga	gagagggttcg	gcgcctcagt
7681	tgccccggag	cagacgtcca	gccccgcgt	gtacccgatt	gcggggcgga	tcgcggcaag
7741	gcggcgcgcg	ccggcagaat	aggcgcgag	gaaggctcag	gcggggcaga	ctgcgtgggg
7801	gaaggaggag	gagagagcag	acggcgagg	agggcagagc	agccgggggg	agggcgagg
7861	ggcgggagga	gacacatgcg	cgctgcgcgc	gccgcgcgcg	ccgccgcagt	ccttagcttc
7921	ccggggacag	gaaaccttca	agaccgagct	gccacggccg	cctccccgcc	cgcctcccat
7981	tctacgcgcc	tgcccacacc	ctcctccctt	ccttccagcg	cctttcggtg	gagcactgcg
8041	gcactcagcc	cgagctgccg	ttttcccttc	gcggggaacg	ctgtgacccc	cccgcaggag
8101	cggcgggggc	gggtgggggg	gccccgggaga	agatggcgac	gcccgggaagc	gaaccccaac

8161	ctttcgtccc	ggcccttttcg	gtagctactc	tgcacccact	tcatcatccc	caccaccacc
8221	accaccacca	tcagcaccac	ggaggaaccg	gcgcccccg	cggggcgggt	ggtggcggcg
8281	gtggcagcgg	gggcttcaac	ctgccttga	accgggtct	ggagcgcgcg	cttgaggagg
8341	cggccaactc	cggggggctg	aacctgagcg	ccaggaaatt	gaaggaattt	ccccgtaccg
8401	cagcccccg	gcacgacctc	tcggacacgg	tgacggcagg	tgagtgaagg	ccgagggggc
8461	ggcaggggtg	tgggtgctgt	ctgggtgtct	gtcgtgcgtt	ccctaaccgc	gtggacagtc
8521	ggagatcttg	tcttgctggg	ggagggaggg	tccatcggcc	cgttgctctc	cgaagaagg
8581	actcgcgtgg	gcgcggaaga	agcgggccct	ggagagggca	cgggggccct	gcctggtccg
8641	gcgatgcagt	gccaggaggg	caggggtgcg	ccgggcctct	gcgcctgaga	gcgaggggtc
8701	tccggctctc	caccctgtgg	ctgccgcgcc	agggaaagta	ccgcggggct	gggacttgca
8761	ggcgcgcgcg	gtgcgcgcag	tgtaaagtgg	aggtttcagg	cggctgtgtc	ggcttcctgg
8821	gccccgcgca	gggctgggaa	ctccagcgcg	gacagcgcg	tcggcgccct	caaccagttc
8881	aagccgtctt	tgcatcgagg	cgtagcccat	cctggtgggg	aaaccagac	aaaggtggca
8941	cgcctcggc	cgagccaggg	gcacgcggcg	agtgcgcgga	gccccctcgg	cgacatcgcc
9001	gagcgatcgg	gcaactcggg	gccgcctgtg	aggtgcccaa	gtttcctctc	cctgcgcttg
9061	tgcggaataa	gagccgccgg	gcttgtagtg	aatcccagtc	ggctgcccgt	tcggaagcga
9121	tcggccctct	tccttcctgt	ccttttgctc	accgcccaca	cagtttgggt	cctaccaccg
9181	aggctacctg	gttctcagtt	acccaaccgt	tgcgggggcc	ccaggaaagg	acggcggtgg
9241	atattaaagg	caaagtaatt	aacgtgagcc	tgttttcctg	tgtgagcccg	gcaagttgta
9301	actcgcgcca	ataagcagca	taactttttc	aagttacgtc	atgtgttaca	tacttctaaa
9361	acgtctgctt	tctcttgggt	agttggagtt	ccaggagcgg	gatctataaa	caggaaagg
9421	ggtgggtgaa	taggattggg	gccttagggg	ctatacagaa	ctgacttggg	ttccggccat
9481	ccggttccgc	tccttggggg	agggtttgtg	tcatagcaag	cgcccaacat	ttccaggggc
9541	agtgggtgct	ccgttttgga	gccgtgccc	agacctcttg	gctgtcatcg	cctcatggcc
9601	cagggctcaga	ccccctgggt	cttgatgggg	ttgactctca	tcacaggagc	accagtttta
9661	atctgttagt	gggcaaggga	gcaggcagtg	ggtttttgga	ggctaggttt	ccagtgctct
9721	ttcccccttt	taattcaata	aacatttata	tagagcctta	atgaatattt	agatacagtt
9781	ctgagttcac	tggaggcaaa	gatggctctc	cagggtcatt	aaactttgta	atgttgattt
9841	agggatagtc	aggaatcggc	aaaaaagagg	gaatttccaa	agttgagctg	tcaacagctg
9901	gggagtgtct	ttcaataaga	cctcttagtt	acatcttctc	aagtggaaact	tattccaaga
9961	attgcctgaa	tcactgactt	aaaaaactac	attaaatttt	acttcccttc	gtctatagag
10021	tgtaccgtga	gctagctatg	ctatggttgg	ctatatctgc	agaaaataga	gctggcatga
10081	ctaataataa	taggtttgtg	tagttcagta	agggggcttc	atagttgttc	tacttcagtt
10141	atgctaagga	ttattttacaa	agtttaggtga	acttgatgtg	tctgctgtgt	tggcatttct
10201	ttgagaaRac	cagtgctttt	agggatcaca	tcttttagct	aagacctcgc	catgtgaagt
10261	ggatcgtgaa	agtgttcttg	ttggcttgggt	tatggcaagc	ttaaagttgtg	agttttgcag
10321	gtttttatatt	atgagctcac	ttgcacatgt	tttatgctca	ggaaaatcta	actggttaaa
10381	attcagaggc	tttatgttaa	ttcctcatcc	aaaaatacag	tattatttgg	gggaagagga
10441	gtggagggat	attcatgttt	gtagaaactt	aagtcaaccc	catcagaaaa	gtatttcattg
10501	cagctagtta	ttttacatga	aattagaata	cccaaagtga	ttaaatattt	gcttctgcaa
10561	aacgtctttg	cctcaggtca	aaagctaggt	gtaaaccaca	gtactctttt	gttctaagga
10621	taaatttaatt	tgagcacagt	ttcttatttg	agagaccaga	aagcaaaata	aagagttaaa
10681	gttacctccg	gtgtacacag	tgttgacaca	gcgtacatgt	aagagttcta	gtacagttga
10741	aagtttgagt	agtaacattt	ataataaaga	tagtggttcc	tccttgactc	atatttactt
10801	taagaatgtc	tcttaaaatt	agaatggcat	atgaagatcc	ttaaaggcct	aagaggcaaa
10861	gtggattttt	ttttccatat	attttctaag	acagggagtg	gtagagcaag	cTTTTtctca
10921	cagttttgtt	gtatattaca	ggccatttgg	tttacgcctt	gtaaaaagtc	aggcttctgg
10981	cagccacctg	tcgagaactg	cacctgggca	atatagccca	tctgaacaga	gtggaaaacc
11041	aagatgtaaa	caaagaaaca	tagaatagat	gtctaagctg	tcactcaagt	gcataatactt
11101	tattgatagg	aaaaaattca	agctctcatt	ttgagcctat	taatttactt	ctatttgaat
11161	caaaataagt	tttgttgtca	gagtgtagtc	ttttaatagg	tgaaccgggg	aaagaaagac
11221	tgatccacgc	aggtgcagtg	agtgcacag	gtgtcactga	tggaaaggga	caaaacagaa
11281	gagccactat	cagcactggg	tcgcctctct	tatgcgttat	ctgcatgata	ttagtcccca
11341	ctaggagccc	agtcataccc	tcattttatac	atgggaagac	tgagacagag	acaagttaca
11401	taatatgctg	aagattatac	ctctactgag	atggtcaacc	ggtgattgga	accctggcag
11461	cctcacgtta	attcatgttc	ttgacccatat	tctcatcaag	cagagggaag	gggttgagag
11521	atagaatctg	ggaggcttcc	tggaaattca	ttcaaccagc	attgattcca	caagaatgta
11581	ctgagccccct	ctgagtgtctg	gccactgatc	taggcactgg	ggacacaaca	gtgacaaaaa
11641	gagaatccct	gtatcccttc	ttgtggagtt	cattctagtt	ggagaataga	gagatacatg
11701	aataaattga	ttttgcagga	taaataggat	gagttacagg	atagatggaa	tatctttgtt
11761	ttataaatac	acagattgct	ttgtattgac	ttcaggtggc	ttgtagcaag	agtataaatg
11821	atgactagta	atacatataa	tctgaagcat	aaaattgggc	caaagaaaaa	caagtaacaca
11881	aattttgcata	tgagtaaagc	aaattttttt	tttttttggg	gctgatttat	taagtagtgt
11941	cgttttgaat	gaagaagttc	aaaagggtat	tttgtacaa	taaaagtga	ttgaaaagga

12001	gtaagaaaag	aaagtggctg	attgcaggag	gttctagaaa	ataaaaaaga	agcccttaga
12061	cttggcatgc	aggacacttg	ttgaaagtct	tgtaaaggaa	agtggggaat	atttcattgc
12121	taattatctt	tttgagttgg	ctctaagcag	gaaagactat	gggccaggca	ctgtggctca
12181	cacctgtaat	cccagcactt	tgggaggcct	agggtgaagg	accacttgac	cccaggagtt
12241	tgagaacggg	ctgggcaata	tagtgagacc	ccatctctac	gaaaaataaa	atattaggcc
12301	aggcacgata	gctcactcct	gtaacgcctg	cactttggga	ggccaagggtg	ggtggatcac
12361	ctgagggtcag	gagtttgaga	ccagcctgac	caacatggta	aaacctcatc	tctactgaaa
12421	atacaaacat	tagccgggca	tggtggcatg	tgcctgtaat	ctcagctact	caggaggctg
12481	aggtgggaga	atcacttgaa	cccaggaggc	ggaggttgca	gtgagctgag	gttgcagtga
12541	gctgaggctg	caccactgca	tgccagcctg	ggcgacagag	agagactcca	tctcaaaaaa
12601	tgaatgaata	aataagttaa	taaaataaaa	aattaggcgg	gtgtggtggc	gcattgctgt
12661	agttccagct	atgcggaagg	ctgagggtggg	aagcccatag	gttgaggctg	cagtgaagac
12721	catgatcgcg	ccactgcact	ccagcctacg	tgacagagtg	agaccctgcc	ttaaacaac
12781	aaacaaacaa	acaaacaaaa	caaagactat	gtgggaaaca	aagagcttac	tcagggaaga
12841	atattgaagt	gcagcattaa	aaatacccat	ttatgtgtga	gggaagagtt	ctgtaggaga
12901	tgcaaagtat	aagacaatac	aggaatacat	gagatgctaa	ataacagttc	caggaagcca
12961	ctgaggccca	aggcagggtca	agggtagggtg	ccaagtagca	ctgatggaca	ttgtgtgtag
13021	aattcagagg	aaaagggtcat	tttcatgagg	atgcttagga	aagactttga	tagcagaggt
13081	aggatttggg	tcagggttttg	taggatggat	aggatttaga	tagaagtgcc	tatgaatatt
13141	ctgtgattgg	aaggatgggtg	ccaagacgtg	gagtaagaac	ctgacagttg	gccgggtgtg
13201	gtgactcaca	cctgtaatcc	cagcatttttg	ggaggctgag	gccgggtggat	cacttgaggc
13261	caggagttca	agaccagcct	ggccaagacg	gtgaaaccct	gtgtctacaa	aaaaaaaaaa
13321	aaaaaaaaaa	aaaagctggg	ttcctgtaat	cccagttact	cgggaggctg	aggcatgaga
13381	attgcttgaa	cctgggagggt	gaaggttgta	gtgagccgag	gttgcccac	tgactccag
13441	cctgggcgac	agaacaagac	tctgtctcaa	acaaacaaac	aaacaaatga	acctgacagt
13501	ctaagattct	gtgaataaaa	gaaacagaga	gggcatttaa	gcctgtcttt	ctggtcatat
13561	tcttctaccc	aactcctgcc	ttataatgga	gtaaaggcta	cagggtctct	atgacccttt
13621	aattgaggac	tccggaaatg	gctgaagcct	aggatttgte	agcctgtggg	aataacaccc
13681	tagctctgtg	tgatctggct	ctctcaaatg	tctgtacaga	ttcaaaggat	cttctctgag
13741	atttgtgggt	ggttttatct	cttgaaggct	gacttcacac	ccctattccc	atcctaagc
13801	aatggagtcc	ttgggatgtt	agcactcttt	agtcgttggg	attcacagaa	attcagataa
13861	caaaatctag	tcattttcaa	caacttaact	ttatttcata	atttttgcca	tcaatatgtt
13921	tcactataaa	tataataaaa	tttttgaaat	cacatgtgtt	tccttaaaat	gtatgttctt
13981	ttatagttat	atccagtttt	gaattatcgt	tctgtatttg	aattaaattt	ggtagttttt
14041	aacatacctg	ttgttttttg	tggtggttac	tatcagataa	atgcctttgt	ttattacatg
14101	tattctaat	attcttccat	tcataaacta	tgattcata	ccttttcttt	octattacta
14161	ttgttttttc	ttccctgtta	aggtagat	aaattccaac	atcaaacaaa	agcttttgga
14221	acaactcaga	caatatctgt	agtgttctg	aatcacgtct	agatgggcc	ttccacatgt
14281	ctgtggtata	tggaacaaac	agtctccaca	aacctacatg	ctatttctcat	tgcttgcat
14341	acaggctgct	ttcttgaaaa	ggcaaggaca	tactaatcat	gcttctagaa	taatagcata
14401	caattatag	ccagggaacc	tctgggaatt	tttgggtctg	gccagtttat	aaaggaaacg
14461	cccaagggtt	gaccttcac	actgagatga	ctctactgga	gagtttgtca	gtcactgcaa
14521	catctgattt	gatttttgg	agggtgttga	tactaggcca	caaattagca	caccctgtca
14581	tttaaaaatt	cttgttttaa	ttattgactg	agactttttt	ttgtgtgcgt	gtgttactgt
14641	aatgaactga	aatctcccat	ggcatggcat	aataactctt	tgtaataaac	attttaaggg
14701	tagatacctt	ttgaaatcca	gtttgacgtc	gattaattca	gaagtcaaac	ttgttttgcc
14761	cgtgctgtcc	atctgggtata	tgtacataga	ctttttctat	tatatcatte	aacttcatca
14821	ttgcttggtt	aaacacaagc	tattttcctc	ttttgtgtat	acatttctgt	tcccccatat
14881	atatagactt	tttatatttc	agctgggcag	catttacttg	ctcttaattg	ctccccataa
14941	ttcccttca	ggtgtggtac	tgcattgtgt	ttttcaaagc	cttgataatt	taaatggcag
15001	aaatgctatt	attaaactcca	tataagggtg	ttatgggttt	gtcactcaRg	ccatgtgtaa
15061	cttgaatttt	tgtctaagcc	tcagaaaaac	ctgttttgga	ggccactttt	Yagaatttct
15121	tgggcttttt	ggaaaattga	acgtctgtct	tggatattgg	cttgagctta	gatattggcaa
15181	ttccaccagt	gattattttca	gtgattgtgt	gaattgttag	aattaagata	atttgggtgt
15241	aattgcagtg	attattttcaa	gttataagat	tRtaactcat	tataaaaYtg	ttaaggatgg
15301	aaggcacaaa	acccaactat	ttttcYRtag	tgaataatct	tttttaatgc	ccaactcata
15361	ggaagcctag	cttctgttat	tgctactagt	cttttggtgc	atgaaagatt	tttgtttttg
15421	catcagaatt	ttgccatgat	atttcagaga	gtaagaaaag	aaaaaaaagaa	aagaaaggtc
15481	agggttccaa	gtcagactta	atgtaggaaa	actaggcatt	ctaataataga	ctcaaacatg
15541	aaatagcct	gtgttattct	tcactcataa	taaagagttt	agtgatattg	cttccatatt
15601	ctgcagtaag	aaatagtaca	tttggaaacc	atctcctccc	ataaagtaat	tcaaaaaataa
15661	agtgaagatt	tttcatgaa	tagaaagttg	atgaactgtg	agaattgggg	cttcccttgc
15721	ttactccagc	catgcaatag	aagggaaatg	cttttatagt	cagatttctg	ttcactgttt
15781	ttcctaacca	tatcctcctt	cactggccat	aaaggaggat	atagctggag	aaaaaaatca

15841	gagaagaaca	gtatgtttga	tttgcataac	agaaatactt	catatttttgc	ttggtaagaa
15901	attattgctt	ctctgttaat	atgatgctgg	tgctgaggca	gctcaactga	ctatacctga
15961	aattcttacc	gaacattttt	aagtaatcgg	ggacagggtta	tttgatttttg	ggatcttccct
16021	agatgtttgt	ttttcttcct	tctgtctgtt	gcttttttcc	ctcagtcctc	ctgctggagg
16081	gggacctaac	acagggtgtg	gatgaaactg	agcaggagcc	ttctagctca	gtgtgtcttg
16141	aatttctttt	ccactgttct	tgccaaatgg	gcagaaagtg	gccccatct	tagtttttct
16201	aatcttcttg	cttctcatgt	ggacttgctt	ttgacagctg	acccctgcag	tggtgtatgt
16261	atagcttttg	cacacgatgt	gttctctgt	tattcttccc	tatttttgat	gtttatgcca
16321	tacagcgtga	ggagtagcat	caatatttct	gatctggttc	agtcagctta	gctctgattt
16381	aactacttta	tctctgcagg	caatgggtgc	aattcaaaga	cagcaccatc	aacacactgc
16441	ttggcttgcc	cttctctcgg	attgccttca	ttttgctttt	ccaagtcatt	tacctcagt
16501	cccttctcct	ccctgttctg	taggattcta	ctgacccccc	tctcatttca	tctgggcccc
16561	cagcctgtca	aaacaaacag	cctccctatc	agccctttct	tgcatagctc	agcactgaat
16621	catacatgac	cttaagccat	ctttgcatta	tttccttatg	gcattttttg	ttttattgccc
16681	ttcctctgat	acttttcccta	tgtttaacct	ttttgctcct	tcactggcct	gtgagatttg
16741	gctagacggg	acttttccctg	atttttgat	ggcttccatg	cccaggacag	aaaatgctag
16801	aaaaacattg	atggattact	ctctttccctc	tttccctacac	ttggtctctg	catttttcat
16861	ccatcatctt	ggatctaccc	tgaagtatac	ccctacttgt	gcatagtttg	ctgatctgct
16921	cagattttcc	tgtgagagct	gatttttctc	acatagcctc	tttttacttc	tcagtgtcat
16981	ccactgtatt	agtctgctag	gggcaccatg	acaatatcc	tcagactggg	tggtcttRac
17041	aacagaaatt	aattttccca	tagctctgga	ggcttgaagt	ccaagatcaa	gggtctggca
17101	ggtttgattc	tcctgaggcc	tctctcttgg	cttacagggtg	gagggtgtct	tcctgggtct
17161	tcacaggccg	ttccttttta	tacacaaatc	cttggtgtcc	cttctctgagc	ccaaatttcc
17221	tcttcttata	aggggtatcag	tgagattgga	ttgggaccca	cccataggac	ctcatgtaat
17281	cttcattact	tcttttaaaga	ctttgtctct	aaatacagtc	acattctgag	gtgctagggg
17341	ttagtgtctt	cattactttt	gggggtcata	attcagccca	taacgcccac	tgactcacc
17401	ctgctgtagc	tggtagctgg	agttcctccc	ccagcctgtg	taaactgatc	ttggacttgt
17461	gtgaggctgc	cctcaatgct	gctcttgcag	gaggcccaag	acaggtctgt	ttgtaacca
17521	tcctgtcctt	ggaatctagc	cctgcacaca	ctagctacac	ctgcatccac	gctgagagtc
17581	ccaataatgc	tcctgagccc	tcgaatgtgc	tgtaaaatgg	taagagagcc	acataatac
17641	ttacaaatgc	atgactcccg	ctcatttgga	cccctcctcc	atctctcata	ctctagaat
17701	cttttcgggtg	gttggtgggt	ccctagagtc	acactttatt	tttagctaag	gtggggtcac
17761	ttctcaccaa	catttgcaag	tgttcgataa	gcaatgttac	atctaaaatc	aatgagttta
17821	tagtgaattc	tgagggagat	gggcagggtg	tttataatat	caaataattgt	ttgaaattag
17881	ccaatgaaaa	actaatcaat	atcttttgag	cacttattgt	gtaaggagca	tagccaaaaa
17941	tataaaggca	tagttccacc	tagtttgaat	tggtatgagga	ttaaagaaat	tagtattttg
18001	aaagtccctt	acacatacag	tactccaatt	aacaagtatt	atttcaccaa	agcctttcag
18061	cccttagcca	aatggaattt	cttccctttt	tttttaagta	tagcttagtg	attaaaaata
18121	gggtttcctg	agcttcaaag	acctgtgttc	aaatcttggc	cttggtgctt	ataaactatg
18181	tcactttggc	cagatgactt	aaacccctatg	catcggttcc	ctcacctgtc	aaatggagat
18241	aataatagta	tgtacctagt	agagtgttca	tgaggacatt	ggttataata	acaacagcta
18301	acattagcaa	gcagttattc	tatgtcaggt	attgtctcaa	gtgctttata	tgcaaatata
18361	tgtattaaat	gtatgtaata	gatgtttcac	ctcatttaat	gcttgctaca	tggtctgtat
18421	taactagagt	gggggtctca	taactccaag	caatgccttt	ctattgcaac	ctgaagaggg
18481	agattagaaa	tgtagaaacc	atccattctg	atctctccct	catttaatat	ctgaggacac
18541	taaggcccag	agagtataat	catttggcca	agggtacaca	gcattggcagt	agggtccaga
18601	gccagggcag	agtctcccag	gctcacatca	aagggtcggt	tctgtggagc	caggaggtga
18661	cattgggatg	aagggaaggc	tttccctcat	ccatagggtt	ctcaggtcaa	acccacccaa
18721	accagccagg	ttgtgggcct	cccaaaaaac	aaaaaaaaat	ggtgtgaaac	ctttaataac
18781	agatgggtcca	aacctgggga	gaataagata	aagctaacgg	tagttacaga	tgagacatgt
18841	tagtgtagga	tttgactgca	tgtggtttga	gggacatgca	caggcctctt	gtcaactgga
18901	ttccacatac	taaaggagtg	tgctgggcag	cccccgactc	acaaacgggg	ccatcagcct
18961	ttcctagcaa	ggctttgttt	cagaacttgg	catgcatttt	cctataaaaa	caaggtggta
19021	acctgtgatt	tggttctccc	gccaacccac	aaagccctgt	ttaacccctag	cagacctggg
19081	ccttcttatt	aatggtgggt	tcttattgtt	tctttggaaa	aatgcaatcc	cagttccaac
19141	atgaaatccc	gaatctggag	aacttggcat	caggctgtga	atgaggcccc	tgctcttga
19201	aaataattct	ctccccctctg	gccattccgg	ccaggggtcg	gaggttcccc	acacctctcc
19261	caggtccctt	gatcagggca	ctcagcggct	gggaaacaag	ggttgccata	agtcagggat
19321	ccacttgga	ctcagggaag	ccactgggac	agggaaggat	aggccgggga	gccacccacc
19381	tgagggagcc	Rtcagggaac	ttagacactt	ctctatttaa	tttttatatt	tcttctgtat
19441	tatctcccca	acttactaat	tttggaattg	gaagcaagtt	agtgaacttt	tctgaacatc
19501	agggtttgtaa	tttggttagaa	gcgaaaatgc	tctttatccc	ttgggtgggt	gtatagattc
19561	aaagggtaat	agatagcaca	ggagcagggc	tcaggtctgg	caaacagtga	gtgactagta
19621	gaaaacctgt	cactttctgt	ccttaatttc	catggggcta	tgtgacatag	tcttcttcag

19681	tcacgtagat	aattggaagc	tgtcagccta	tcattagcat	agaggctaaa	gcacaaactg
19741	tggtctgggtc	tttgtccctg	ctctgccagg	acatggctac	tacttcactt	ccctgagtc
19801	cagtttcctc	gtcataaaat	ggggataata	attatagaat	tattgtgagg	aacaaatgaa
19861	tacatatgtc	aagtgccttag	gtcagtgcat	ggcactgatt	ttgccattat	gattattctg
19921	tatcttttaga	aaagagagtc	agtgttataa	ggcacagaaa	gactaaattc	ctagaaatgc
19981	tgtgacatta	ataaaataac	atcagtatcc	agcaagcttg	ggtgtacatg	ctttcagtg
20041	gtgtacatgt	tatttatatc	ctttcggcag	tcttgggagg	atggtagcat	aggtgggtatt
20101	cattattcca	tattgcatgg	gggaaacag	gctggatgg	gaaatgactt	tccttcgggt
20161	atgctgctag	tcagcagttg	agacaagacc	acatcttctg	ggttttcatg	ctgtttttat
20221	tcctcaattc	aacggtgttc	ttaggtttga	tttaatttct	gctggggccg	ggtgcaatgg
20281	ctcatacctg	taatcctagc	aatttgggag	gccaaggtgg	gtaaatcact	tgaggccagg
20341	agtttgagac	cagcctggcc	aatatggtga	gacccgtctc	tactaaaaat	acaaaaatta
20401	gctgggtgtg	gtggtgggca	cctgtaatcc	aagctactca	ggaggctagg	gcaagagaat
20461	cgcttgaaact	gggaggcgga	ggttgcagtg	agctgagatc	acgccattgc	actccagcct
20521	gggtgacaga	atgagactct	gtctccaaaa	taataataat	aataataata	ataataataa
20581	taataatgac	aacaataatt	tctactggga	cagaggagtc	ttgaggggac	tgagtctcgc
20641	agagaagagt	gtgttcaggt	ttctcacagt	agagaatggg	gaaccctatt	cctgtgccac
20701	gttgtgtatt	caggtgaagg	gtgctggggc	cagctcttgt	tttggcctcc	agattggagg
20761	ggtgaggaag	tggtggctgg	ggcctcagag	gagggcctgg	ggtgcctgtc	ttcagtcctc
20821	acggaagtct	tgctattgat	ttaaaagctc	ttccaagtta	aaaaacaagc	aaacaaataa
20881	aaaaccattg	acagtcctct	catgttcttg	gcaaagaaca	atttatgaat	aattttatgt
20941	tcttgacatg	ctctgaagat	gaatagcaaa	cacttctacc	agaagtgaag	aatggctac
21001	taggaggctg	agcttttagga	ataaatatat	agtagattca	attgttatct	ttttttccc
21061	tcactctttt	actggcatcc	agcatttatt	ttgaacttca	gtcagaagtt	atttcactaa
21121	gaacctgtgc	atagcttcaa	catctccctg	aggctactga	tatggtaact	cggagggggg
21181	aataatgttt	gtagggccag	actcatgtct	gggtccaggg	tctttcagaa	tgttcctctt
21241	tcagtttctca	gaacaccctt	gtgaggtcgg	gagtatcatc	gttactgaga	agggcactta
21301	tcacaagtta	ggcagtcaga	gtatatgtgt	aagccagtag	taagacagag	ccgaaattgg
21361	agcccaaaca	gttgttgaag	aagtgcagga	tcctcctcta	ttttagtgtc	actgaatatt
21421	tagagctaga	ggaagagacc	agagaaggag	atthttgcaga	tgagaatcaa	gcctatggca
21481	tttctacaga	tggtcctcgg	ggtgacatag	agatttttaga	gttcaccccc	agctatggca
21541	cggatttctg	actaatgtga	tttccccac	tgatctgaaa	gtgagctcag	agaccccgga
21601	tatatgcatg	gggagttcca	gccctaatac	tcacttgttt	gtgatggcca	gctatcagct
21661	ccattttctaa	gaattgtgtt	ttaaagcaag	tcacaattac	tRgtttttaa	aattcttcaa
21721	ggtatctgtt	attaaaaaat	atagacacag	atthgggttt	tcagctttgg	tataaacttt
21781	tcttggttaga	ctttgtaaca	catgacctaa	ttgaacagtt	tactgatatt	ttctataaac
21841	agtttactga	tatttacttt	ccctccagtg	agaaatgaac	gaatttcac	gagtgcaga
21901	ccaagtttac	tttctagaac	attacctata	gaatacaaat	gtttacattc	caaagactgt
21961	tttattatgt	aataagaatc	atcaatttga	catataaaaa	tatgttttat	tttgtaatat
22021	agatctttct	atacattaat	tacctctaca	atthttataga	tgaaataggg	cagcacaaga
22081	ttatggctaa	aaaacagatt	ttttttgagt	gcatgcaaat	caaatccaac	ttaaaggaat
22141	ctgtgcaatt	gacggcaaat	agatctctct	ttgtgcagtt	cccttgagg	tgaaaatgat
22201	aatctctgca	tgcaattttaa	ggcaattttg	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg
22261	tgttttctctg	gtgtttttaga	aagcctggat	aaatccaagt	taatttttagt	tgtagcagtc
22321	attccatcgg	ttataggctt	gtaatcaact	cttagctgtg	gtttgttcag	atatttaca
22381	tgaatgcttt	cttgtatcta	tacagtactg	taaaagggtac	tgtgtctata	Stttataaag
22441	acaaatgtgc	ctaaatctga	tatattttat	aaaattcagt	tggcagctgt	gaaggcttgc
22501	ttttgtctct	tcctcccctg	cacaggctaa	gtccttggaa	tcagtatgtt	aatgaagact
22561	tggtatctca	atatgaaatt	tgacccatgg	gagttctgag	ataatgcgtc	ctgatagagc
22621	ccctaagtgc	tgccctcttag	gagggcgagt	gacaaaagaca	gtataactca	gctaagtcct
22681	ttcataatca	ggcagccttg	gtctttctca	agtgtttcca	tagtttctgt	gtgtgtgctt
22741	tcttgtgttt	ctgtgtctca	gacctcttct	ctccttgggc	agtctgtctg	taatctttta
22801	atactgcatg	gtagatttga	ttggatgagc	cccatcacct	ttggataaga	aaaggctttt
22861	ggacacaagt	ccatgcaaca	Ytcttttctt	tagagaacac	aatagatcac	tggaacatc
22921	ataagaggtg	ctctcagagg	gtgcttataa	gtggcaccat	cctaaatcca	ggtttaccgg
22981	ggacagtcgt	ggttttataac	tcactctggc	ctaatagaata	atagaccctt	cttttagcac
23041	agcatgattt	agataagtta	tatggtcacc	ccatgcacat	gggatgtgg	agataacacc
23101	ttttaagaat	aaagtgaat	ggattcaagt	gaagtgaact	acctaccat	gtctagttgt
23161	gtcagattga	atccagaatt	ggtccaggtt	atcctgggtca	tttcaccagc	tacctatatt
23221	gctgtccttt	gaaggaatta	caatgcttgg	ctggttctct	tttttaactt	tcagtccaag
23281	aaaccctctt	agtaaagtct	tagacatttg	ttgtttat	tagttgttaa	aatacagcat
23341	taattttaaaa	attacaaaat	attgcataca	taggaaggag	tccgtaccct	atgtagacat
23401	gattttaagta	atgatgataa	aacagacata	gtgacagaag	catatagaaa	ggtgatcacc
23461	tggaaggagg	tgacagagag	atggaaacat	tctatgtctt	agaatggtga	gtacataggt

23521	agggtgatgt	atttgtcaga	actcagtga	ctctgcactt	aaaatgggtg	cattttactg
23581	tatgtaaatc	ataccttgat	acagtagact	aaaaatttac	ctattttacca	ctatccaact
23641	taaaaaggag	aacattttcca	ataccttttt	agcttttctgc	ctgggtgccga	ttcctgactg
23701	acaagctggt	ttaaaacaca	gaataactga	ggcttttattg	ctaaatcaat	gcctcctatt
23761	agaagactga	tattttctggt	ttagtggacg	ttctttaatc	tttatttttcc	ttatttcgaa
23821	acgtcttggt	aacctctttt	acgtagtctt	actcataaaa	tatttctaatt	ctgtattttaa
23881	caagcggcat	tttatttttag	tgatagaaaa	aactgtggct	gaaccaagtt	ttctgggcatc
23941	actggaatga	tgttagctaa	tggagtccaa	tgttgaaagc	atttgttgcg	cacctactgg
24001	atgtcagaca	ctttgcagac	acgatcttat	ttctcctctg	agcctaataa	taggtatcat
24061	ttcccttaaa	tttcctcatg	tgtagaattg	aggcctatag	atgttaagaa	cttgcccaag
24121	atccccatgc	aaaagtaatt	aggagaatta	ggagtcaggg	tttgagcccc	agggatctaa
24181	gtacattggt	aatgtagaga	actgcccttt	gctgggagaa	gttaagcact	tggctggctc
24241	ccctgggtgt	atcgtgtagc	accaggacca	tgtggcgtat	cacttccaaa	tttccctaaat
24301	cactgtgatc	ctgcatactc	cttgcccaga	ccccctctgt	ggctctctcc	tgcttgccca
24361	ccgaggctca	gctgcctagc	agcagatcca	cagggcctct	ggtatctggc	gccaaaccgc
24421	ttatcctgct	gggttttctt	ctagctcccc	tcaaggccct	gtgtggctcc	agctgccttt
24481	tccattagtg	ttcccatatg	ttccaggtta	aggcctccgc	tttgccttcg	attgtacctt
24541	ttctgtcttc	attgcacttg	ccaaataaaa	taaacaaaac	aaaacaatga	cagaccccag
24601	ctctccagtc	cctgtcaagc	agtcactcca	aaattttaat	tcctgcctct	tctaggatgc
24661	cccaaatgcc	agcctccctc	tctgcaactc	tactcatctt	ttcatctaaa	acagagttta
24721	taatgattga	gttggctctt	aattgttctt	taattccatg	tcttttttct	ctaataagaa
24781	taagaatgat	aatagcttca	tttattgagt	cagatattgt	gttctccact	ttatgtggct
24841	tcctcagct	aactcccaga	gcaatcttgg	gaggtaagta	cacattttct	tcccatctga
24901	cagactggaa	aattgaggct	cagagagaaa	tgcctgagat	cacatggcta	ctaagtggca
24961	agactgggat	tcggtttgac	ttgagtttga	ctactctttg	ctgacttctt	gcaaatgcct
25021	tgaggagagt	agaaaagaaa	ggaactgcac	attgactgtg	cctttgctat	ggcttgcat
25081	cgagatgggt	tagcttatta	tctcacaggg	tccttatcct	gtgctaagaa	tcagatctcc
25141	aaaaaaattg	agaattagat	ctctaaaaaa	cggaggctct	gtttacataa	ttcattcaag
25201	gtcacagcgc	tgattagtgg	agagataaaa	attaaaccag	gtctttctgc	ccctagagat
25261	gtgccactgc	gtgcatgagg	acactttatt	tctttttctt	actctctgca	tggcctagca
25321	tattagggaa	tatttaggga	acttttcaat	tagttatagg	ttgcttaaaa	aacaacaaaa
25381	acaaaagtcc	ctttttattct	gttagaagca	tgagcatatc	acacatattc	cggtagggaa
25441	gagtagtacc	ataaagaatg	ttggattttt	agaagagaca	acgtaaaagc	aaaggtgagg
25501	tgagatagac	acttgaacag	aatggcagtc	acagaatttc	ttatcttttg	agagatttac
25561	ttgagggatt	ctacctttac	tatagtagag	gagtgtctaa	aattaaaaga	caggagaaga
25621	tgacttagta	ttcccatcca	gatttttaaa	ttgctgtgtg	gttagctgtc	tccagtacag
25681	cagcttctac	gtcctcttct	ttaactgcct	ttctcaggca	tagaagtaaa	aagaagtaaa
25741	aggtaacggc	cagttgccct	cggcctctcc	attgggtggat	ttgcagagtg	ctgaaattca
25801	ccctagagtt	tacccttcaa	agtaactagg	cattaaaaaa	ctcatacctc	caaaaaagga
25861	actcctttgt	gttccaacaa	ttgttacaat	atgtagact	tttgggtgat	aaaacacacc
25921	tccattaact	gtccagcat	cttgctgcatc	tcagctgttt	gtgagagtca	tccagcatga
25981	gggaggatgg	gtctgcatca	ggtgcagttc	ctgggccttg	gaatatgtcc	tccacttttg
26041	tttgatcagc	ggcgtgtttc	cattgagtg	gctgtgtgcc	aagagaacag	ttgtcagtg
26101	tattcgcagt	gatgcacatt	agcatgagtg	ccagagattc	aagtggcgac	gagtaccaac
26161	tctgccttcc	acctattgct	ttttgtgctc	tcagaaactg	gagaaacgct	agctttgcac
26221	actgggtttt	ctcgaagcgg	gctaaggcta	tcatttccaa	gggctgtgga	gcttcttgga
26281	actgtgcatg	ctgtgaacac	tctccagaaa	tgatgaaatg	ccaaagtgtc	caagtcacca
26341	tgcccttggt	cactggctcc	catgcctttc	ttgctgtgac	catccatgag	gcttagaacc
26401	atcacagctg	tttaggaaat	ggtcttgga	tccaacagac	caaggtcaaa	ttctggctct
26461	gtcacttagc	agttgggtga	cttagggcaa	ttcatgtacc	tctctgaact	tctcatttct
26521	catctgtaaa	gtaagtga	gtaataaac	ttatgggtt	gctgtgatga	ctaaaataac
26581	aaattctaa	tatgtggcac	ttggggatg	tgtgtgtgag	agagagagag	aaagagcggg
26641	agggaggag	ggagagagag	agagagacag	gcaggcagtg	gctgtttttt	tgaacggcga
26701	tgcccacaga	gggggagtta	tttttgagga	ggagggggtc	ttctcactgc	tcccaattct
26761	catccatgct	ggggacgtgc	atctgtctct	gccactactc	tggaacata	tctcttgcg
26821	gatacagctg	tctttctggt	gttcccttgg	tctgccagca	cagggtgag	gccattccct
26881	gcactcctct	tcctacttgt	acccccactg	caggccatgc	ctggggactc	cctgaccctg
26941	ctgtccttct	cagcttgga	cttgctgctc	tttatgggga	ggacttttgt	tctctttgcc
27001	aggctgcctg	agaaacaggc	ccaggtgggtg	tgaagtaata	atccccagac	caaatatgtt
27061	tatcgtctta	tttgagccat	cctgagactt	actgcaacac	tacttttaagc	ttagtgaatt
27121	taatttggat	cacagcctat	tatacatatg	gcattggtgc	agcttttagt	aaaagataaa
27181	ataggtgcgt	gaaatgaaag	agacatagat	gccagttgga	gaacagagtg	atctaaaagc
27241	aaattctgca	tagcaatatg	aatacttcag	agtggacagg	ggtatagttt	gaaagaagct
27301	acctataacg	actgttttca	ttatatgtac	taaagaaaaa	aaaagctcga	caaaatctcc

27361	ctaactcgtg	gagtcatatg	gaagatagaa	taagaagaca	gtgcacaatt	gtgggcaaaa
27421	gaaggtaatg	aaacattgca	aagactttta	aaagcttttt	ttaagttcta	tattcacagc
27481	cacatttctg	aaacttaaga	ggtgctgcat	tttttttcat	tcattcatcc	atccagcaaa
27541	tattttattat	ctagtgtgag	taaacttgcc	catgagccca	tccctttcca	ctgcctgta
27601	ctgtatgtga	caaggatagt	agcgtgctca	gtagcttgca	accacgttgc	ctgggtgagc
27661	gactcctcca	tctactggcc	cagtgcctt	gggcaagttc	cataacctct	ctatgctcta
27721	gtttcctcat	ctgtaaaaga	gaaWtgataa	tagcacctgt	cacagagggc	tttggggaat
27781	tgaaaatgca	aggagtgcag	agcaatgcct	ggcatctgag	agacattcaa	tacatgttaa
27841	ctcttatcat	ttgttttctg	acttattcat	ttgcttcttt	attgtcactc	tctctttcca
27901	tcattaaacg	ctgagagaaa	ccttgttttc	tcttttctca	tgtaactact	atccttggtt
27961	cctaggacag	tgtctgccc	tggtagcagc	ttagtaaata	tttgagccga	gtaagtgaat
28021	aaatggcttg	atgaatgaat	ggaatgcag	aaggctggac	tctgtgctgg	gtgcttccta
28081	gacaggagt	aacaaaactt	ggtgcctgct	gggtatatgc	ccaaaggaaa	gaaaatcagt
28141	acagcaaaga	gatattctgca	ccccatggt	tggtgcagca	ctggtgacaa	tagccaagat
28201	ttgaaagcat	cctaagtgtc	catcaacaga	tgaatggata	aagaaaatgt	ggtacatatg
28261	cacaatggag	tactattcag	ccataaaaag	aatgagatcc	tgctatttgc	aaccacatgg
28321	atagaactgg	acatcattat	attaagtga	ataagccggg	cacagaaaga	caaacctcgc
28381	atgttctcac	ttatttgtag	gagctaaaaa	ttaaaacagt	tgaactcatg	gacatagaca
28441	tagagagtag	aatgatgggt	accagaggat	gggaagtgg	gggatgggaa	ggaggtggg
28501	atggttaata	ggtatgaaaa	atagaatgaa	tgaaggaa	agatctagta	tatgatagca
28561	caacagggtg	actatagtca	aatatatata	tatatatata	tatttttttt	ttttttgaga
28621	tgaggtctta	ctctattccc	cagactggag	cgcagtggca	tgatctcagc	tactgcaac
28681	ctccacctgc	ctcgttcaag	cctcccaagt	agctgggatt	acaggcacac	accaccatgc
28741	ccggctaatt	tttgtatttt	tagtagagac	ggggtttcgc	catgttggcc	aggctgggtc
28801	cgaaactcct	acctcacctc	aggtgatccg	cctgcctcgg	cctcccaaag	tgctggaatt
28861	acaggccact	gtgcctgacc	atatagtcaa	taataattta	attgtaattt	aaatataatt
28921	ttacagtaac	ctaaagagta	taattatcct	ttgtaacaca	aaggataaat	gcttgaaatg
28981	atggatacc	atttatcctg	atgtgattat	tgcatgcctg	tatgaaagta	tctcatgtac
29041	cccataaata	tatataccta	ctatgtaccc	ccagaaatta	aaaattaaaa	aaaattaaaca
29101	ttaaaaata	caaaacaagt	ggtgttgctc	gcattttaga	gcttccagct	agtcaagtga
29161	aggagaagtc	tagagatcta	taataaccaa	agggtgtaag	tgtcaatgca	agggagaagc
29221	ggggacacca	tgctgtatgg	catcaaggaa	gaagtggcat	tttccgacta	tggtgagtgg
29281	tggtactgata	aacaagagca	ctttttttct	ttatccatta	gaagacttaa	aataattatt
29341	tgtagatccc	acttccccac	tttttaaagc	atttccttct	caattattcc	ctctcttcaa
29401	acatacaaac	ttatacgcat	attctaaaac	ttgtaaaagg	aagttaaaca	cattttaatc
29461	atttaaaagg	ttttttaaaa	agtcaaccca	tgtaacttaa	aaatgattat	gaagcacttg
29521	tttactttcg	gttgtctggc	atttatttgt	gtgtgggggt	tttttttttt	tttttttttt
29581	ttggtgtatg	ggctattttat	tgctttttct	taattttttc	cttgtgtgtt	tggaaagcaa
29641	aatccaaaca	aagtcctttt	tggcagtgg	aataataaat	attctgttta	gcctttggat
29701	gaggtttaca	tgatgtatct	ataagcacat	tctggaatat	tggaaataaca	ggttccaatc
29761	cttttgatca	atggatagt	ggataaatga	aaagttccaa	ccaaaatata	cctactggac
29821	aagaaccatc	atctttttaa	aggcagaaat	cctctttagg	gtccagcctt	cctttcatte
29881	ctgttctaca	ttgtcaatga	ctctgaccac	aggaacagag	aactgctgga	taggggtgag
29941	tgctgctggt	cccctctatc	atcaccagca	caaataattg	tattaatttt	tttttttagga
30001	tggtgagttta	agctcagctg	tgcaccccag	acattgatga	agatagtcta	tcgtatgcta
30061	tagagtcaca	gaaccacaag	atgcgtgtgt	gattatctaa	ataaagagct	cttactgcaa
30121	ccttgtgact	ctgggtgatag	aactgggcgt	gggagggcct	tctcattcat	gtgctgtgca
30181	cacaaatgat	tctctaagtt	gaaggacaca	ttgctctgca	gcataactga	aaaatctgtg
30241	catgaatgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	ttgtgggggg	gatgtggtct
30301	cattcctact	tgcaaaaatc	ctccagaagg	ggttttcaaa	gcattgcaca	gctgttgctt
30361	gtcagcttct	tgtgttgact	agatgtcagc	aacacttctc	caactttgga	caatgacttg
30421	aaagcaagat	cagtgtgtgc	tcatgggatg	gatttaaagt	gaattgaatg	atacagaaac
30481	cattgcctaa	tttatattta	ggtaagtcac	gttttctttc	tctctctctt	taaatttttc
30541	tcgaaatgca	tgtatgaagc	tctttttggc	taaaatgaga	gaatgaagca	ggctcctgag
30601	ctcttgagt	aagaggatcc	attgagaagt	ccagagggag	agagaatgag	cttgattaac
30661	ggcagcacag	gattggcagg	agtgtatata	aatagtacat	catctcagag	tgacctataa
30721	ctcgactttg	gaatgcattt	ttatttttgg	cagtggattt	acctccctgt	gtcctgcctg
30781	gacaaacgtt	ggaattagtc	agcattcctt	aagcacatgc	tgattatggc	caaagaaagt
30841	ggcaggcttg	aggttgctg	gcgaccagga	gaaagccagc	tggattagac	atttgccgtg
30901	agccttttgt	tcatgtaact	tatggggcag	tctcctaagc	cagagaaaaca	agaaactcaa
30961	aggcagtttt	agcaaggttc	tcatgcatgg	cagcatattt	taataactaat	gctcccttta
31021	ggcaggccct	aaataagaat	gtcttggtct	aatttttttc	cgtagactta	agattgttgg
31081	gaaaaaactt	gttagaattc	ccaagtga	ccatttaaat	gccttggtt	ctgataactg
31141	ggagtcctga	gtgaagttgt	cagttgctgc	tgaactcagc	agaaataatt	gcgctcccta

31201	tgccaatgat	ctgtgcatat	tttgagacca	ttctgagacc	aaggaggcta	tcacaacatg
31261	ttcacattgc	tatcgttgat	tagagccctg	agctttacat	gttggtctcct	gtgggtgtcct
31321	tcggaatatt	ggagagccat	gcctctttcc	tttcaatggt	ctcaacctgg	cccagccagg
31381	attagaggct	cccttctgag	gttcagtaga	attttatatt	taccctttat	cattattcac
31441	cttactgtat	tgcactttctc	tgtctacagt	ctccagcatt	ccattacagt	gctaggacta
31501	tgttggcatt	tgtattcggt	tttcattact	tcataagaaa	ttgccccaaa	tttggtggct
31561	taaaatgatg	cacatttatg	atggcacagt	ttctatggat	caggagtcca	gacacagttt
31621	ggctggattc	tctgccctag	gtctcacaag	gctgtactca	aggaactggc	tggggcttct
31681	ctcatctgaa	gcttgaggatc	cttttccaag	ctttcgagggt	tggtggcagg	attcgggtccc
31741	ttgtgattgt	ggaactcatg	gtggcctttt	ttctttcagg	ccagcaggaa	aagggtctctg
31801	acttgagggt	gggccccagt	ctttctttta	ttttattatt	attttttgag	acagagtttc
31861	gttcttgggt	cccaggctga	gtgcaacctc	gggtccatct	cggctcactg	caacctccac
31921	ctccagaggt	caaagtatcc	tcccacctca	gcctcccaag	tagctaggat	tacaggcaca
31981	cgccaccatg	cctggctaata	ttttgtatct	tttagtagag	atgggatttc	accatgttgg
32041	ccaggctgggt	ctcgaactcc	tgacctcaag	tgatccaccc	acctcgccct	cccaaagtgc
32101	tggtgattaca	ggtatgagcc	actgcacctg	gccccaggtt	cttcttttaa	aggtcatcac
32161	taattaggtc	aggccactc	aacagttgat	taaccttaata	tatatattgca	aaatcccttt
32221	gccattttaag	atgacacaat	aattggaatg	atatcccatc	atattcacag	gtcctacctta
32281	tactcaaaaag	gtggtagggt	atcacagggtg	gatatagtgt	tgaagaagg	aatgaatgaa
32341	atgaatgaat	tctgatagct	ggaagcacag	caactagctt	ttttttagc	atgagtgaga
32401	gaagtattgt	agttccttat	tgaagaatgc	agttgcaaca	gtttatcact	agtgacttgc
32461	ttcttggtaa	ctataagaca	taagcacaga	agaaacagca	agaaagggaag	gaaggagaaa
32521	agagaacgag	aaggaaatca	ttattttgga	ccgatttccg	gtgRtggttt	caagaagttc
32581	tccccaccct	ccagactgct	gtcattttaca	attgctagag	aacaagagct	gggagcctct
32641	ccttggttct	gctcctgtgg	catagccatg	ggggatgcag	aaggggcagt	gtgccccMta
32701	aagggaatcc	ccgtcagctt	tggttttagt	ttccagagat	tgagcctctg	ccttagcttg
32761	agagaaaaac	ttagtgtgag	tctttatttg	ctcccttata	atcacaggca	catgtggcac
32821	ctcctcccc	atctctgccc	cacctcttcc	aacagtgaga	tgcattaggt	gtggaaagga
32881	agccttcatt	tcaaaatctc	ctggctgctg	ttagctttaa	aaatcaaatg	gaactgtgggt
32941	gctctgtggt	taattttgga	gagcagaagg	gatgcgtgct	ttgtggcgcc	tgcatagtga
33001	aagtctccat	gacttaatac	cactgcccag	ccttagaagc	ttcccatgca	gacaatggcc
33061	ctgcttcccc	atcagttctt	gggcttgggt	gccaccttat	ataMacattg	ctgcttttag
33121	tgcacgcgag	aagcaaaaag	atgacacatt	aatagcagct	ggaggatctc	attaaaattc
33181	Rgaggaaaat	atgaaggctg	gagcctggta	atgtacagac	agtaagtctt	tgtctgtctg
33241	attaacaatt	gcatgYattc	taagtgcctt	gctggatggt	aagaaccagc	atttcagcca
33301	gaagtaccat	gttcttctcc	ttaagaaata	cctatttctg	tttttcaatt	cccttatccc
33361	ctctaaaaat	aatatgctgg	tcttttggac	tcccttgaga	actgaaacag	cttagcaact
33421	ggggcagtc	ctgagagctg	gtgcagaaac	aactgttgct	agcttggtat	tgaattatgc
33481	agtcagatgc	ttacatcgtg	ggagggggat	gtacgggggt	tgtggccttc	agatgcttac
33541	atYgtgggag	ggggatgtac	aggggtgtgtg	gccttcacgc	cagaaaggac	aggggaagac
33601	tccctggttg	gcgggaggag	gggagatgct	gggggttagg	cccttacctt	tatttatgtt
33661	gttcaactggt	acagccagac	tgcaacctct	caactcttca	gacaagacct	aaggccgaca
33721	aaattcactc	cctcatgaag	aacattgggc	tgtttctccc	tggtgtgtag	gtggtgggag
33781	tgaagtagag	gggctcgctg	tgtgcaggga	aagggtgca	ttctcaagaa	tttgtatgta
33841	gcccttggtg	taaaaaacaa	atattccaaa	ggaatgcttc	ttttggaacc	ctgggtccagt
33901	cctcgtcta	agcaggagtg	cccttatatg	cttgactgtg	gatggcatct	gattacgggc
33961	agaagggttg	actggattgc	tgccctgatgc	ggacccacag	aacacagttt	aactctggct
34021	ctgtcagcag	tcatcagtg	ttatttttcc	aactcttcag	agagatatta	gtcgttatga
34081	aaagtgtacc	atctcagaac	tccatggaat	aagctgccaa	tataagacag	gttgaaatta
34141	ttgtgtcttt	ttgcctagag	ttatttttag	ttatttttct	ttccttgga	tttcttttct
34201	tttttttttt	taaccatata	tgctaacttt	gtagctgggtg	gcaaaaaaag	attattttcc
34261	ctatatatta	gacttcttca	gctgaattga	ggaatagggt	taccagctct	tttcttggag
34321	tttaacactc	tcaatctgat	aatatgaata	acacctcaat	aaccacaaca	agagcggatc
34381	atcttcagta	tcttttctga	attacctttt	ggcttttcac	caccttggtt	gtccgtttca
34441	gagtttgtct	ttctgtaaaa	tcatgggaag	gcacttgatt	actttgttct	tgatgtacct
34501	gagtaatgcy	aacatctgaa	taattgaaac	aaaatcttgt	ttaaacaaaa	cttttgttct
34561	agaagcatct	gtcaaattgt	tagtcatctg	acaattttaa	atgtcaaagc	aagagttggc
34621	tgatcttaac	cttagccttg	tcacattatt	aacttcatgg	agcttcagag	taacttgaga
34681	agacttgcag	tctgtacctg	aaaaagccag	gtttagaaag	tggctcatga	agaagtaaaa
34741	cccctcagtc	aagtggaaatg	tgcaagggaa	cccctaggag	cattttggca	tgaacatttc
34801	ttttgagttt	ctagcccaag	acctctttat	aacatttggt	cctctcccag	gtcctcctcc
34861	tactgcctct	aattttttaa	taagttctgt	ttaaactcatt	tttgataaat	aagctagctc
34921	tgagaaagct	tggtagaact	attttctactg	aaggagtctt	aaaactgatc	atatctcttg
34981	tcatcaaatt	ataacctgac	tgttgccctg	tatctgaaat	tctaagatag	tgaagagagt

35041	tatttttactt	tttctttag	ggcccttctg	tcattacttc	cttgctagaa	tcctatgtgt
35101	gataatgagg	gaatgaaatt	atccaaagg	taatttgcag	tatgcatcct	accagcagt
35161	tgaaactact	gcagcgaaac	tgcaaggaaa	cctaaccat	gttttcttgt	aatgctaccc
35221	tttgggggtc	atgcatgctg	ctttctgctc	caacagtcct	aacgtgggaa	ttggggacac
35281	ttactttttg	acctggaaca	aatctctaag	actcgtaagc	caccaattga	ttcacctttt
35341	gtagcagggc	ttgtatgaca	aactcttttg	tgagaacaaa	aggaaagtcc	cttccctttt
35401	tgccatttct	gttctgtaga	tttaattagca	aagacgtagt	ggagaaatgt	gtcacttgca
35461	agacatttct	ggcacggagt	ggaccaagt	gaaatgctta	aatgtgttgc	agggatatca
35521	aagggactga	gcttttggaa	caaagaggaa	cagagcccaa	aggacaRttg	gtacaaatta
35581	cacaattttc	agttggttct	aataaagtcc	taggaaaagg	ggcttatgca	caaaggctga
35641	gctcttgcc	ctgttggttct	tcaaattccaa	aatgtcatta	aaatactagt	agtggttgtc
35701	atattttta	aacatttttg	atttgaagt	tgtagctgaa	acttactctt	tgtaattgga
35761	tgtatcactg	ggtgagcttc	tagactgcct	caccaatgga	tctagcatta	ccttgggttaa
35821	tgaaataagg	taaaacacag	tcctctgtgta	ccagaaagca	ggttcattgt	tcagctgcat
35881	catgtcattt	ggaactttgc	agggtggtt	ttcctttgct	ccttcaatct	cctcatttct
35941	ttcattgccc	ggatgaagcc	cctgcaacag	gagagctttt	gcctcctctg	cctcacctca
36001	gccccctaga	cctctgtgtg	caactggagt	atggtgttaa	gtgagcctgg	agcagtctta
36061	ggaaaaggct	tcaagtgagg	ctaggtggga	aggcctggga	cagaaataaa	ttcaaaacaa
36121	caaccaaaaga	actgataggt	ggaataaaag	aagtgggaca	gaaaagattc	aaaggtggga
36181	attgaaagag	agaggtacag	taagaatagc	tggcgctccag	tgaacatgta	cagtggacca
36241	tgcacagttc	caagcacgtt	cgatttatta	tttcatctaa	gactcacgag	aactctataa
36301	gacaaaggct	tgacattctc	atccccgtgt	tagatgtgag	gaatctggag	cacaaatgtg
36361	taagtaattt	ttctgttgag	aaatggacct	atcaggggtt	gaacccaagc	agtttgcctc
36421	cagttaatgt	ccttcgccac	tttatattga	actgagaaaa	ggcgggacaa	aagtcacaac
36481	caaagggtgc	tttggccttc	agaatatctc	acctaactgc	atcaatatct	caccttctca
36541	agcccacagc	ctcctgtttc	tttttttctg	ccccatttcc	ataattaatc	atccttgcaa
36601	taaaagagac	ttttggaaat	tattttggga	aaatgaagtc	tttgggaggc	caaggtgggt
36661	ggatcacttg	aggtcaggag	ttttaggcca	ccctggccaa	tatggcgaaa	ccccctttgt
36721	actaaaaatt	caaaaattag	ccgggcgtgc	tgggtgtgtg	ctgtagtccc	agctacttgg
36781	gaggctgagg	caggagaatc	gcttgaacct	gggaggcgga	gtttgcagt	agccgaaatc
36841	atgccacagc	actccagcct	ggcgcgagca	gagagactcc	atctcaaaaa	aaaaaataaa
36901	aagaaagaaa	gaaaatgaag	ggtatagctg	aaatctgctc	atttcatgat	cttgacaatt
36961	taaaaaacct	gattgctgtt	gcagtaggac	tgtttaaaaa	aggaaagaga	aaaatatatt
37021	tttaaaaaatc	tgagcatggt	gaaagacaca	aaggaccttt	gttacaagtc	actcctacca
37081	atggaagaac	aaccgttgct	ctcatctttt	tttttttttt	tttttttttg	tcattttcatc
37141	ttccagctgg	ctgagatcaa	taagatagtc	tctatggttg	cttttgagtt	cctttgtagt
37201	tcttaatgaa	gggatgaaca	tattgttttt	tttttttgac	acagggtctc	attctgtcac
37261	tcctgtctgg	agtgcagtga	cacaatctta	gctcactgta	gccttgacct	cccggtctca
37321	ggtgatcctc	ccaccttagc	ctcctgagtg	gctgggacca	caggcacgca	ccacctgccc
37381	cagctttttt	tttttttttt	ttttttttgt	atttttagta	gagacggggt	ctcaacatgt
37441	tgcccaggct	ggtcttgaac	ttctcagatc	aagccatctg	cttgccctcg	cctcccagag
37501	tgctgggatt	ataggatttc	gccgctgtgc	ccagccacat	attgcttttg	atgtagttaa
37561	acagtttaact	gctaccaagc	ttctcacaga	ggttactggg	ctcaggaata	aaaaggcttc
37621	tactccaaat	ctggtagcac	cttgaccac	aaagacttgc	tctctcagta	agaaaaacaa
37681	tggcagttct	atatatagtc	aagagtggag	gcaggaaaag	actacatttc	tgaaaaatgt
37741	gttcccaaga	aatccatggt	aatcaatgtc	agagaccatc	cagtttttat	atattgcagt
37801	ctcaagaaat	attctgccac	agattctcta	aggactgatc	ataaaacaca	aaaccttttt
37861	atctttttgta	agtttgtcag	ttctaacatt	aaatgcttaa	ttttgacagt	tacaaaaaca
37921	agtgcattct	agcttttctt	ttttaaagtt	tcaggtgaaa	tttgctataa	tctcatgtcc
37981	atcctcattc	aggcatgctg	aaagcactaa	tattctgttt	gcttttgagt	gctaagaaaa
38041	tatagcccat	gagattactg	gccaaattga	aaagaatgag	gcttattgta	ccaagatac
38101	tttcttgata	aattaagtca	atgatctatt	tcctgcacta	gaggtcttga	ttgggtatgt
38161	catagcgtaa	ttttgtttgc	cttctgaggg	acacctattc	tcttgctctt	taacttggaa
38221	agctcctaac	tcagaagcca	ctctgtgaac	atatggctcc	aaagattgct	tcagattcct
38281	ggctgcctgt	gctaagtcca	gtactcacct	ttggacagga	tatctaattt	ttagattcta
38341	gacagttgaa	agtaactaac	ctctctaggt	ttttcttctg	tttccctagag	ttaagtttat
38401	caatgtgatt	taatctgact	agattactat	acatttttga	acgtaatcat	ttatttgtgt
38461	ttatattatt	tgctcttgtt	tcagctacgg	aatttttcta	tttttttttt	ttttgagaca
38521	gtgtgtcact	gtgtccccca	ggctggagtg	cagtggcgca	atctccgctc	actgcaacct
38581	cgctgcctca	ggttcaagca	attctcctgc	ctcagcctcc	caagtagctg	ggattacagg
38641	cgtgagccac	tgcgccagc	caaggaaattt	tgcttttctt	agtcagctag	attttaaact
38701	ccacgaggtc	aggaaactta	cttttcacac	gtttttcctt	cttatagaat	aatattttct
38761	cgatatgttt	tagaggtggt	attttatagg	cagtgggttaa	agactgaaaa	ctttaaccac
38821	tttaaaactct	gaaattctag	gcaatgggat	ttgtactgaa	agacatagga	taattatgac


```

38881  acttcaatta tagtcggttc aattcaccac tggggtggtt ctaagttagg ttcattggtg
38941  ctttctccca ggaaatctaa agtactttac atatatagct ccattgtttc tcatagcatg
39001  caaccttaac ccagtgtttt cacatgtgat ttgagagctc gataccctaa ttatgggtcca
39061  tgcaccagca gcaccggcat cacctaggag ctgggtggaa atgcaggctc tcaggcctca
39121  cccaaggcct catgaatcag agtctgtatt ttaacaagat ttcgggtgat ctgtgtgcac
39181  attacagttt gaagagataa gacaccaaaag ttaactattt ttcctgttat cacatggcaa
39241  gtagatctcc agtcctgagt ctgggttgca tagcttgtga tgttgatgtt gcaaagatgc
39301  aggtgccttc ttttcccatg atgatgcaga gccccctcag ctctgttggt agcctgtgat
39361  ggtctctaca ctccgtgttt ttcttcttag gctctttatc tgtacctgct ttgcttttgt
39421  cccttaaggt tttgtccctt tgatttcctc ttaagtgtc tgaacttaac catttttgtt
39481  ttcttaattt taacatggcc atttcagtct gatgacagtt ttataatcaa aaccacatg
39541  caatcgatta ataatgtttt ataaagcaac agttgctacg aaaaaaagtg gatcttctct
39601  agatgaggct atagttaaaa tatggatgga atttgtatgc caattgggaa ctgatggaga
39661  aatgtatttt ggtctgacat aaaagttcct ttctcaaact actttcatat aactaagcaa
39721  tgaaagagca gagaattctc aacttgcttc actactggaa agtaaaggga aatttatggg
39781  ggtaatagtg acattgagaa gtagttcttt aaagaactgc atttttcata ggtaactgga
39841  aacatttcca gctttgaaaa tgaacagata tctttttctc tcatgtttag taaggcttta
39901  gaaatgagta ggaacaattt ggtagaattt cttacctgta tttcaaaaca tgctgagtct
39961  acttggttaag taagatgttt taaagcatt tggctccttg aattctgcta ctcataagaa
40021  tgcatttcta gagccgggct tgggtgctca cgctgtaat cccagcactt tgggaggccg
40081  aggcRggcgg atcacgaggt caggagatca agaccattct cgctaacacg gtgaaaccag
40141  gtctctacta aaaatacaaa aaattagcca ggcatagtgg tgggcgcctg tagtcccagc
40201  tactcgtgaa tctgggaggc ggagctttca gtgagcagag atcatgccac tgcactccag
40261  cctgggcgac agagcgcagac tgcattctca aaaaaaaaaa aaaaaaacaa aaaaacgcat
40321  ttctataacc tgcaccgatt ttgtagaatt tagtgtgatc tgaagctagt tgcctctgaat
40381  atgctttatg taaatatgtg tggctcactc aagtaatttt cattctttag attacatctc
40441  tcgatctttt tggatgtaac tggatttgct ggagttaaga aWgaacgttg tcattccccc
40501  tgcccacccc caccttcctc ggctattgat gtaataaatg taattacggt actaagaaat
40561  cctttccagc tgaaggaaca aaccacaaag tctgtgcttt aaatccttgt aattttttct
40621  ttccgtttta aatctaagcc atatattcag atttgggcgt gctttttgca ggtgtctatt
40681  tttagagctt tccctgtttt gtaatttccc agttctcagt catgagtatt ggggaaaaat
40741  gcacattaaa attgggtttt acttttgttt ttttttaaaa tttgttttga gttttataac
40801  tttaaaagcc tcgtgcattc ttttcacagc ttccagcccc aggaacatgg caatggaaca
40861  tgtttatttc cagtctagca tctagaaaaa gctaactcct ggatgacaca aagtaaatga
40921  ttgcaaaaag aaagagagag aggaagaagag gatgacaggc aggggaaaga tcagggggct
40981  tgaccacaga tggggcgcat ttactgaaat taagcatgat agttgttgcc caggaaagtc
41041  tcacagaggg ccacatgtgg ggtcagaggt tgcctgtcac atgtggttg cagccagaaa
41101  cagccatagc aaggctgagg ggagcatcat tcattttttc attcagttag gtttttgttt
41161  ttgttttttg agacagggtc tccctctgtc actcaggctg gagtgcagag gtgcaatctt
41221  agtccactgc agcttccact tcccagggtc caagggatcc tcccactca gcctcccga
41281  ctacggcatt tataggcaca cgccaccaac acccaactga ttgttgattt tttagttag
41341  atgggggttc gccatgttgc ccaggctggt ctcaaactcc tgagctcaaa gcaatccgct
41401  ggccctcagc tctcaaagtg ctgggattac aggcgtgagc caccatgccc agcttcagtt
41461  aatgtttatt gagttccctc cctgggttaa ggaatgatac ttcagcttta acagtgaggg
41521  gtaggcgcca ggatcaggtc ctgggatgtg tggctgtatt tttctggcat gagcagctt
41581  ctacggcatt gttcttttc aagcaggaata gtttattaga attttgaca gatgcggttc
41641  atttcacca ggattttgtc ttagttttag cagcagagac agtgaatctc actgggtctt
41701  aatcaggaag aggagttggc cgttaaatga cctcaggagt ggcttctggc ttagaccctg
41761  gagatgccgt gtggtttgtc aaggatcaac tagctcatga caaatactag aagagactta
41821  ataaatctgg acctttcttt accaatccaa gttgcttgag ttgtaaatga tgtaaaaaa
41881  gcaacctaga gacactagtt gttcagggtg aaatctctta tttccattag ccctgattct
41941  agagaagagt ggagagttgg ccttaggtgc cctctttgtc tttgactgta tatgtggcac
42001  actttcttca gccaatggaa tggcatatcgt tgtatgccat ctgcaagagc ctgataagtt
42061  gcattacaaa tagagtgtgc aaacaatagt actgattaag tgacaaatgt tgaggcctgg
42121  gaaactgatt ttggcactga caaattaagg cagattagac cgcctttcag gacactttt
42181  ataatgctac gtgtatgtga aagaagaagg tatcagaaaa aacttaatag agtttcttag
42241  caagagtact ggaacataat tgtggatgcc taagagggaag tgtttgagtc aggaaaattt
42301  cagttgtcct gaattgatga acgagtctct ttttagtaaa tgcttctttt aacggctcat
42361  gttagtttag acttagagac aaattacaga gtagcttatt gttatttctg ttgttgata
42421  ataaactttt ggagatgaac ttcattaaat gccctgggtc tgggtttgtt ggtgaaaagc
42481  tgaataaat ctattgttgt taacattctgt ggtgatatag acttgaaatt ataacttag
42541  catgggttag aggagctgtg agaaacagtg aaaattcata tagcacttat tatacacagt
42601  gtatgaaatt aataatctct gcacctcct cccaattgct gaaattatc ttacggtttc
42661  catttcattg ttgtaccata gagtatgtaa gggatgctgg ggctaagact attagcatta

```

42721	tggaatatt	taaatataat	ttaatatgag	aaaaatcaga	ggaatgtgca	aactctaggc
42781	tgtatttctc	cctaattgggt	aaatccagtc	atatatctct	gacagaattg	taaaaacact
42841	gagttatttg	atctttcctg	actgacttat	cttgagtgca	tttatattta	taacatgctg
42901	tgcaccaaag	catgaaaagc	agaagcatat	agtttcgctc	tcatctgaag	taataaaata
42961	ttttttattt	acatagtcaa	tctgggatag	attatagtag	aaaacccta	caaactctgag
43021	atactgaaag	tggtagccgt	tttcagagat	aaataagtaa	cctatttagc	tgattagatg
43081	tctttgctac	tcacatggta	aataaggcgt	aactttgctt	gtttactcct	ttaaatttcc
43141	ttatagatgt	acttactttc	tttttatatg	tttaaggctc	tagtagaatt	tcatcccttc
43201	ctccagcttt	ttattgtata	aaatgttaaa	tgtggaaaaa	ttgaagaatt	tatgtagcat
43261	agcacatctg	tatatctact	acctaaatta	attaattaac	atggtgctgt	atttcgtttc
43321	tctctctctg	tctttttttt	tttttttttt	tttgagacag	tcatctcgct	ctgtcaccca
43381	ggctggaatg	cagtggcgcc	atctccgttc	actgcaaccc	ccgcctctcg	ggttcaagtg
43441	attctcctgc	cttagcctcc	cgagtagctg	ggattacagg	cgcccgccac	tacatccagc
43501	taactttttt	gtatatattg	tagagacggg	gttttgccat	gttgccagg	ctggtctcga
43561	actcctgatc	tcagggtgatc	tgccctgcctt	ggcctcccaa	agtgtggga	ttataggtgt
43621	gagcaaccgt	gcccagcctg	catttgcttt	ctctatttgt	gtgtgtatat	gtgtatgtga
43681	aggtatgcgg	atgtgtgtat	atatatatat	atgttttttg	ctgaacagtt	tgagattaat
43741	cagggtgaag	agatgtcatg	atgcttacc	tctctctcta	tatatTTTTT	aaggcagggt
43801	ctcactatgt	caccagcgt	ggagggcagt	ggtgcaatca	tagctcactg	cagcctcgaa
43861	cttctgggct	caactaatcc	tcccacctca	gcctcccaa	tagctggaac	tataggcaca
43921	tgccaccaca	ctggctaata	tttaatttta	tttctataga	gagaaatctt	accatgttgc
43981	ccaggctggt	ctcagactca	agtgatcctc	ctgcctcagc	ctcccaagta	ttcagattat
44041	aggcatgagc	cactgtgcct	ggcctaagtt	cttgagcatc	tgtttcttaa	tagtaagggc
44101	attctcgtac	ataaagacac	taacatttca	aaatatgaaa	aatgaacgat	aattttctca
44161	ttatatctgg	tatccagtc	atatttaaat	ttatctagtt	gtctccaaat	tgtcttttat
44221	ggcgattttt	tcttctaaaa	ttcagatata	cagtgtctcat	gcattgcaat	tcatatgtga
44281	gtcctaaact	cttttatggt	taatctagca	tctagaacag	ttctctcact	tttttttttt
44341	tttttttttt	tttgagatgg	cgttttgctc	ttgtcaccca	ggctggaggg	caatggtgct
44401	atattggctc	actgtaacct	ccgcctcctg	ggttcaagtg	attctcctgt	ctcagccttc
44461	cgagtagtg	ggattacttg	tgcccaccaa	cacgcccagc	taatttttgt	atttttagta
44521	gagacagggt	tttaccacat	tggttagcct	ggctctgaac	tcctgacctc	aggtagtcca
44581	atggccttgg	cttcacaaag	tgctgggatt	acaagtgtaa	gccactgtgc	ctgggatttt
44641	tttttttttt	ttttgacatt	gactttagaa	agagaccagg	ccacttgcct	tggaggacat
44701	cccacatgga	ggatgtgtct	gattgtttcc	tggtgaggct	gttcagtttc	tctcattcat
44761	gatttctctg	aaactggaag	ttaggccaga	agtttgaggt	ttattagttc	agttaatctt
44821	ttttgggaag	atgaggtaga	ggagattaa	aaaaagtcag	tgctatcttc	gcagatgtat
44881	ttttcttaaa	accctggcct	ctgtgatatt	gagaaagtca	caatatggga	aggtagacact
44941	tgtcaacttc	tgattagagt	ctcttctcta	ttttaaaaga	catgaactaa	aagttgactg
45001	ggtttctctg	cctccccctt	tcacctatc	tctttctatc	tcctcttctt	tgctcttcca
45061	tgccatagag	tatgtgccac	taataggaca	ttttcttaca	atgtggcaat	ttccttgtat
45121	acgtttatgt	ctagcccacc	attctatgag	ctgctgacag	cagggaggtg	ttgctttttt
45181	tgtgatttct	tgaaacccaa	tataattcag	tgttatcaca	ggccaattta	ttaacattta
45241	aaactctctt	cttggccagg	cggtgtggct	cacacctgta	atctcagccc	tttaggaggc
45301	taaggcagga	ggatcacttg	agcccaggag	tttgaggctg	cagtgaacca	tgattgcagc
45361	actacactgc	agcctggata	acagagcaag	accctatctc	aaaaataaaa	taaaaataata
45421	aataaataaa	agtctcatct	taaatttgat	ggggaaggga	attttctgga	tcacaaactg
45481	aattgtctat	gcttggcagg	cagatgtttt	ctcaaatggg	gaatggattc	accttctttc
45541	tcttccctag	tttagctact	ttcactccct	tctagtcttc	atatattccc	aagtgaagtg
45601	agaagatagg	tctgtatatc	agtgaactgt	cttttctttt	Wtcttttcga	tttctctttt
45661	aggagctgag	atgtttcttc	tctgaacttt	atttttcatt	gcttatggct	tttttttttc
45721	cccctgtaag	ttctctgcga	gctatacacc	tgtggaata	tagggactct	ccttcttttt
45781	caaagagctc	agaatatcgg	tgagacaaag	tcccttctctg	cctggacatt	cagattggcc
45841	ctatggactt	aaattgtcaa	gcaagtatac	tttttttgtgt	gccccgcaac	ttggccataa
45901	ctttgctata	aatggcaaa	ttactgaatt	gccttggtctc	cataccaaat	gttggaataa
45961	tttttaggtg	ataaaaacac	atcttctata	tgaaagtatt	tcttttaggg	tcccttgata
46021	agcatatata	tgtgttactc	attttcttaa	tgaatttagYt	cctttcttca	attgataagg
46081	taaaccctca	ttaatatcta	atttgttgga	atgaaaaatga	ttccagtaac	atttttatga
46141	cttcaatgtg	actcccttat	gaaaaggcat	aatggaatgg	ctactttaat	caaatttgaa
46201	gaaaaaagaa	atattattct	gaatacgcac	ttcataaata	ggaaaggcag	ttttcagcat
46261	ttacttttag	tatcatttat	tcagtggagag	cttgagggtg	acccccaggt	tcttgaggc
46321	aattaagacy	aaagtcatga	ttttggagag	ctctgggttt	agtaaaacca	cattagatta
46381	gagttcttag	accaggaagg	ggccctagaa	aaccatcaga	tcaagcctct	tgcccttagg
46441	caaataggat	actaaccctt	gttgacagac	aagcatattt	gaaaccaata	atggtgtttt
46501	gaacaaaaga	aaaatagggg	cttgaggcct	gggttttcat	cttttattgc	actcttctca

46561	ctgtaattat	agtttttattt	ttcaattcaa	ttcattggat	ccgagaggag	ctcattttatc
46621	gatttgcaac	ctaagtatta	gccatggagt	cattcttgaa	ctttgccgtg	gacttggttg
46681	aaaaatgttt	ctggttttca	aggcttccct	agtatcatat	ggtattggaa	tcaaaaatag
46741	agtattgttt	taaaactgtc	agcataggca	gaacttctca	tccagcaacg	cattccgtag
46801	ctcaagtcct	cattttgaac	ataggattgc	aatgctgagc	ccctaccatc	tttccactgg
46861	aaggaacagt	tgggcagctt	tttaaggtag	ggcgccctcg	ggaggactcc	attaatgagg
46921	aatcatgggtc	tcctgacttt	ggattgatca	ttttgttca	aaacaggagt	attcctgagt
46981	cagagtcagc	cacagcttat	aggcaacaat	tttcaggaga	caacatccca	agttaactgc
47041	ttatggcccc	caaaactgta	aggcttgtga	tagtgcccca	gtgaaatatg	ccactgtggc
47101	attcccattcc	tttttagctcc	tggattttctc	cttttcagcc	cataataggc	aagggcaagt
47161	aagggaaacta	acaattatat	gagcacttcc	tctgtcccac	acatggtgct	aagtccttta
47221	aagaaaacaa	ggtttttcttc	actggttctc	tatggccatt	attagagaat	gctggagtca
47281	ctgatctgga	aaaacctgtt	taattccttg	gtctcaattt	cttatctgta	aatgagatt
47341	aaattggtaa	aatatgtgaa	gcacttaaaa	agtatgtttt	aaatggtact	tagtgttacc
47401	cacttcatat	acatttctgc	cagctctatc	atacctaate	actcatttga	tcatgtcatc
47461	ctcttcccag	tctctgttct	ttgagctgaa	gaatgtttgg	tgattccctg	ctgcttcttg
47521	gcaaaactcc	ttagtttggt	gtgtgatctg	gccccgtctt	acctctttag	actgatagat
47581	cccttcattc	tcgcattggt	attaaaaatc	aatatagacc	atgattttgg	gtttatggct
47641	tctgtaagtt	tgtgagggtt	ctctacctaa	aattccccta	ccacctgtcc	tcctatgtcc
47701	cttccctcta	cagtctccac	ctgttcaaat	ccccatgtcc	cttcaggaca	cgggtcccag
47761	cctctgcccc	aatgtagcct	ttcctgtctg	ccttttaaact	gggagttaata	tcctcctttg
47821	gattactagc	cctgtattta	tagctgtttt	gaaaccctat	ggaagtgtag	gctttgaatc
47881	atcaaaaagt	attaagtgat	gttcagtact	tccatttact	cagtactaaa	gttcatcagt
47941	gtattttcaa	acatttttca	acaagacttg	aaataaaaaa	aaagtttttc	cttaagaggt
48001	tttttcttct	tttcttctct	tttttttttt	tttttttgte	catttgacat	ctgaatcctg
48061	aattgactag	acaaatttgg	tttttctagt	cagtgtttaa	ctgggacatg	ccattcttca
48121	aacatttcta	ggaatcccaa	tacctagtag	ctgattcggg	cgtgctggag	aatacaagg
48181	cagtaatcaa	agagcctaca	cagagagaca	ctgaatttta	gaaccaggat	aatcaaaagt
48241	gacttttagtg	aaacgtcacc	acgatctgac	gtgatctgaa	taaaccacaa	tctcagagag
48301	tgaggatatg	ttttgagtaa	attgtttctg	tgtgttggcg	agagtggctg	cctagttaac
48361	cggtggcagg	tctgtgagtt	tgcggtctgc	cccttcggtc	cagctgcttg	ctgataggct
48421	ggcccaggcg	ggatccatcc	ttctgatcgc	caggctcctga	tgaggctggt	gccaccacta
48481	catccattcc	agggagactc	ccaRtctctg	tcagtttctt	ctgcttttct	gaattctaaa
48541	ctcactctga	tctgatattca	ttattttgtct	gttgtggtag	ctttggtgaa	gttgaccac
48601	aaataatgat	atagaagaaa	aatgaactt	tttttcttct	ttcctgtgtc	cttttatcag
48661	gtatcatttc	ctctataaaa	ctaattttta	gttgatagag	tcttaggtct	atagccactg
48721	ttgaatgcac	ctaatacaggc	catctccttg	aactagagaa	tgtttgcatc	ataggataga
48781	taccaggttc	cctgagaggt	gggtaccagg	tgcctgggaa	gtgaagtaac	ttgccaagaa
48841	cagagagcca	gggagtgaac	ttgcaggcat	ttaagcctag	gcagtctgag	tcccgaaaac
48901	aagtcgaagc	atttattgKc	ttccattttc	ttaccatgt	tcttactgtt	ttgcttggac
48961	tcctctttta	atctcagtg	aaaaaataaa	aaaagtacat	atttaacaac	tgaagaaaat
49021	aaacagagac	tgtagtaaat	ttccaagcta	atagcaagtt	atgtaaaaaa	tactacttgc
49081	tgatgaggtt	tgtaagaacc	tcctagaacc	ttatagaaca	agtgtggaat	acttgttttc
49141	tgttaagggc	tattgacccc	aagaaaattc	aatgaagcga	tctttgcaaa	atagtaaatg
49201	atttttcttt	gtctccctgc	agaagcaaa	aggcttagac	ctctagccac	tgttgaatgc
49261	accaaatcag	gccacctcct	tgagctacac	agaatgtttg	cattatagga	taggtgcaa
49321	gtgccctgga	aggttggtgg	agagagagat	gacttccagc	tgagattgta	cagtagaatt
49381	taatatattaa	agtttctcga	tttgacgggg	ctgaattagt	tcatatgagt	tcataagtag
49441	gaactgcttc	gcttaatttt	ggtgaataaa	atatcctggc	tgcaaaatag	aaaaaagaaa
49501	gtcaccctca	ttttcaaattg	tatggttagt	gctttccata	gaaaacatat	ggctaaatgt
49561	gtgtgttttt	ttaatattagc	tctgaaatgt	gaagactgta	ataagattcta	gtaacaagag
49621	cgcagtttag	aaaaacctga	ttgactctag	tttatgtaac	tatacaggct	atataaagaa
49681	agtctcataa	atRaacttca	tctaaagagt	gtggcagcta	taaccttctc	caactttcag
49741	gctctgggtg	ttctcctagc	ttcttctgag	tttatagtct	ttatgtaatt	attaatacca
49801	tgatgatcat	tctgggggtt	gctttttgcc	ctggctactt	taggtttaat	tttatccat
49861	tcctgggtcat	ttatYagtttt	gttctgtcat	ctccccatac	ccatcgagaa	ccctatttct
49921	ctgtggcaga	ctttaatgaa	tgtaaaccga	gattatgtgt	tttctttctt	tgggtgaactt
49981	agaggatgca	taggccagggt	gccaaagtgt	agaagaaacY	atacctttcc	tgatgggtgt
50041	gcattggctag	cctaccccga	ccatgcagta	agtgaacttc	tcaccgtgta	tgtggcagac
50101	ctgctctgca	tctccacagg	gtttctgggt	tcttttccag	gctgtttcca	tattgcacac
50161	tcagactggg	taggaaagca	ggcaggggcac	agttttatat	gtacgtcagc	attttcactt
50221	gggtgatggat	ataagctaac	ttttgtagct	ctcctagggt	tattactgac	atctgttccc
50281	ccctgagcat	atccagctgg	ggctgaagcc	acatctgcac	tttaaaacttc	catctacctt
50341	tatgtttttca	atgtaatttt	agattttcatt	tgtttttaat	tgtgaactca	gaagagatga

50401	gttcagctg	tgactcagtg	tgagttccta	tgtcataatt	ctacatttct	ggagtgccta
50461	cacctgacag	gaataaactc	aaggaaacaa	aaagaacagg	aaaggcctta	ggcacatact
50521	ggttgttcta	gatgtggagt	tctgctggca	actgccacag	agtgaatttg	gccacatgga
50581	atccaaagca	cggcctcaca	ttgttttttt	tttttNgaat	attttatgta	agacagggat
50641	gttaaaggta	gaggaagtaa	ttaagaatgg	cagctatgtc	tttaaaatta	aacacatcgg
50701	tcataacttt	aatgaaggYg	taaaagtgtt	tttagtttta	aacaggcaaa	aaggctttta
50761	aatacaagta	accagttttg	agactttaaa	aagcagaagt	ttttcatgcc	agtgtctcct
50821	attttagttt	caaaggaaag	gaggaggagc	tgaggtttgg	atgggttatca	tagatgaggg
50881	agttgacatg	atcaaaaaatg	tttttttccc	tgggaacaca	tcttgagtgc	ttatctcttc
50941	taatagataa	agggtcgggtg	aattttgaat	gtttcctgca	gctctgaaga	aacactgtga
51001	tcctaattgaa	caccgaggaa	agcttgat	gcagccctaa	atattacctg	cttcaaggag
51061	gcagcatggt	ttgggtacagt	ctgatcatga	ctataaatca	aagcatcttt	acttctccag
51121	ggagataaaa	aaaatcatgt	gttactttat	aaggatcttg	tagttgcagt	atgtctgtca
51181	gatgtttcca	tttttatgat	ttaagacact	tggtgctgct	atgaatagca	aattggaaaa
51241	attgggcatt	ttttaaattt	tgaattttat	cttagcatat	atctggaaat	gaaatagcga
51301	tcttggaaca	gagacatctt	tgttagaata	tgaaaagaatg	tcagtgaatt	ctaacttgaa
51361	gctacattga	gatgacatgt	taaaggcatg	aatagacaaa	gggaatgatt	ttcaggaagt
51421	gccttctgga	gactgtggga	aatcccgta	tgggtagagg	aacagcttgc	gattggatca
51481	aagtcacgca	aagttacgca	ggtgtagat	gctggtagtc	aacatggcta	ggtccatggc
51541	aatcgatcat	ttgcccagca	ttttgcctgc	tgaatttggg	ggagtgaaga	ggacattttc
51601	accaccccag	gaagatagta	actagctgct	tatgtatgta	gctcagagcc	cttgaatttt
51661	caagtgcact	tttaaaagtt	tccccagaa	aggtttgaat	agacttccta	tctattactc
51721	taagtttagg	tctaagaaaa	ttcccaaag	ggtaataca	gtcgatcctc	aatctttgag
51781	aattctgtat	ctgtgaaatt	ggcctacttg	ctaaaatgta	tttgtaagcc	ccaaatccat
51841	ccgtgaggac	tctttgtggg	cattgtgtga	tatgtacaga	gcagcaaaaa	aatttgagta
51901	gctcaaggta	cacgtacccg	gtggaggttg	aacaagggtga	cactctgcct	tcttgcctca
51961	gttctcatat	tgtaaacgag	tgtcctggat	gcagacaatt	cagtgccatg	tcttctgcat
52021	ttttgggctt	tttcttgggtg	atttcaactgt	ttaaaatggc	tcccaaaata	tagggctaaa
52081	gaactgtcca	gtaggacttt	tttttttttt	agatggagtc	ttactcttgt	cgcccaggct
52141	ggagtgcagt	ggcttgatct	tggtcactgt	caacttccgc	ctcccgggtt	caagcaattc
52201	tctgctctca	gcctctcaag	tagctggggg	tacaggcatc	agccaccacg	cccgctaatt
52261	ttttatatatt	ttagtagaga	cagggtttca	ccatgtcgaa	caggctggac	ttgaactcct
52321	gacctcaggt	gatccaccag	ccttggcctc	ccaaagtgtc	gggattacag	gcgtgaacca
52381	ccgcgcctgg	ccctgtccag	tgtttttaag	tgcaagaagg	ctgtgatgtg	ccttataggg
52441	aaaaatatac	gtgttagcta	agctttcttc	aggcatgggc	tttagtgctg	ttggctgtga
52501	tttttagtgtc	aatgaaccaa	caacttagat	caaataagcc	attttttttt	taaaaaacag
52561	aaacacacat	aaaatgacgt	tataagttgg	ttgatgaaaa	ttttgtgacc	agaggcttgc
52621	aggaacctaa	ccctgtgttt	ccccagaaag	ctaaggatta	gtatttccta	tggcggtgtt
52681	caaggcaact	tcatagacag	aactaccatc	atgaataaca	aggatcactg	tgtgggctga
52741	gtgcggtggc	tcatgcctgt	aatcccagca	ctttgggagg	ccgaggcagg	cagatcacct
52801	gaggaaggga	gttcaagact	agcctggcca	acatgatgaa	accctgtctc	tactaaaaat
52861	acaaaaatta	gccggtcaag	gtggcgggca	cccaataaaa	tgacaataga	gtttaaactc
52921	catggtttct	gaggcatttt	ctgagtaaat	tggcatacag	cgtatgtact	ctttcctcta
52981	gaagttccag	aaacaacact	atttctttat	gtgcaaaatg	gcctcttttg	agcagccctg
53041	gggcagtttt	gtctggccct	cttgacgcca	gggtgcccc	gtttagtgtg	caattgggat
53101	aaaaataggc	aacacaggaa	cttgctgtgc	tccggggaaa	agacgcttgc	agattttatg
53161	aaattttaca	tttgatgca	tgatattctg	taggttcaag	aaaaaacaat	tcaatttcaa
53221	gataacattc	tacagggtaa	ataaaaattta	atttcaataa	atttaagggg	aaaagttgtc
53281	catctttcta	tactttctgt	aacttttgta	gtttcaccac	taacaacaac	aaaaaaaaaa
53341	gaacacaaac	aaaatagcct	tgctctgggt	tttgaggaaa	tggttttgca	aggctatttg
53401	gttagacaat	gaattagagt	cagaacttcc	gggatgggct	ttcggttaagg	gaaatgccta
53461	ggctgctgca	aagcctggat	tcaacttaca	caggatcctt	gagaagtgtg	tcttcgcata
53521	cagaaccatg	ggcaatgctc	tatggtataa	aaaccccgaa	ggtaaacact	gtctgatata
53581	tattttttat	aattgcaaaa	tacacataaa	cttaccatct	tgactgtttt	taagtgtgca
53641	gttacatatt	caattgtcgt	cacagtctcc	agagctcttt	tcactctgca	aaactgaaac
53701	tctgtacat	taaatgactc	ttcatctccc	tctctccagc	tcagccccgt	gaaaccacca
53761	ttctgttttc	catctctatg	aatttgacta	ccctaggtac	cttatataaa	tagaagcata
53821	gagtatttgt	cttttttagta	ttggctgatt	ttacttagca	ttatataatg	tccttaaggt
53881	ccatccatac	tgtagcaggt	gtcagaattt	ccttccctct	taaggctgat	taatatttta
53941	ttgcatgtat	acaccacatt	ttgtttatcc	attcttctgt	caatgaacat	ttgggtgtgt
54001	tccacaattt	gactattgtg	aataatgctg	ccatgaacgt	gggtttgcaa	atatctcttt
54061	gagaccctgt	tttcaatttt	tttttgttgt	tgtatactca	gaagtagaat	tgtgtgatca
54121	gacggtaact	ttatttttaa	tttttttgag	gatctgccat	aatgttttcc	atgggtgttc
54181	caccattttta	cattcccacc	cacagtacac	aggggttcca	gtttctccac	atccttgcca

54241	acatttggtta	ctttcattttt	ttttggttagc	tgactgataa	ttatgactaa	ataatattgt
54301	tgaagaacta	ttacaatgtc	aagaaatttt	ggccatcagt	gatagtctta	tgattaaact
54361	tagtagtatt	ttattattaa	acttagtagc	atttattagt	agtagtttta	tttcagaaat
54421	atttgcatth	tccatgtttc	tagccctca	attatgtagg	tagaaacaaa	taatatagaa
54481	tcaattttacc	ttatgtttacc	ttagaactgt	ggccacagcc	tagtaggtgc	tcagttcatt
54541	tttggttaaat	gaatgaatca	atgaccatga	agacagttca	ggttatttgt	tatggagata
54601	cgtaatggga	ttggaaaaca	tgtagggtaa	atgtatttagc	cctctcctgt	gttaactcct
54661	tagccctttc	aaaactaaag	gtattttggcc	atgtgcggtg	gctcacgcca	ccgatcttaa
54721	ttttttttatt	ttaaaaatag	cttacttaaa	aaatagcttt	gttgtattaa	aaaatagcac
54781	ccagctaatt	tttgtatttt	cagtagagac	agggtttcac	catgttggcc	aggctggctc
54841	tgaactcctg	gcctcaggtg	atccaccgc	ctcaggctc	ccatagtgtc	gggattacag
54901	gcgtgagcca	ccacacccgg	ccaggctttt	taacatagta	acataaacat	ttttattctc
54961	acaatgtcct	tatggtttgg	gaatggtggc	ttcctgtcct	tcagagtctg	gtgttatttg
55021	tttttttgggt	gtgtgccttg	cagacacctg	cacttgaaaat	ctttcaggta	ttttgcagtc
55081	gttttctcag	atggctggat	tatttcaagc	caagaataac	agagttaggg	tcaagactgt
55141	gaaccgtatg	gtcagtcctt	ctaagggaagt	atttattttt	attcactttt	gtttgccttt
55201	ccttgggtca	tggtcctaatt	tgtgttccc	tttggctgca	gttattcaaa	actgaattta
55261	ctgctgagcc	taagacagtg	tttttcaaac	ttttaaaatt	aagacacca	gtaagaaata
55321	tgtttttacag	ttttttttaca	tcttgaccag	atttacacac	acacacacac	acacacacac
55381	acacacacac	agaacctgaa	gtattttgtg	caagttgcag	atgtttacag	ttactgctaa
55441	ggatattgtg	ttccatttaa	ttacataaaa	atactgttca	ttgaattgat	ttttgactc
55501	acattatttc	Rtcatgcagt	ttaacaaata	ctgagcactg	gtgtcacaga	actttgtcgt
55561	tcattgtcct	tagcaacagt	cagcttttct	gctcatgact	gatggcctat	tgctatgatt
55621	ctgttagcgc	tttaaagcaa	ttgtattgtc	aaagtcatgt	tagctctggt	tgtgtattca
55681	gttactcagt	ttcaccttta	caggcagctg	ctccttggga	aatggggctt	gctgagcagg
55741	ttgaatgttc	catagaatca	gatctatact	ttggggaact	cagcagtatg	ggaatcacia
55801	gccaacccac	cctatccgtt	aaagggtcct	ggccatttgg	tctccattgt	cacttagact
55861	agcacaacat	cacctacctc	atggatgctt	tgagggcagg	gccggttttg	tttttctttt
55921	tattcctcag	agcccgcagc	atgattcctt	gtgtggagaa	tatgcatcag	ccttggctct
55981	ggctcctctc	tcactgtccc	tcttctcctc	cctacagaaa	acagcaaagc	atgtatagaa
56041	agagatgtgc	agactgtgtg	gaatgccttg	aggggttaact	tgcttagagg	tgtagagact
56101	ttgtgtgtaa	aaacttatgt	ttgaggcaaa	ttgcaggaaa	aagtctctta	aatgaaggaa
56161	agataagatt	ataagttttt	taaaagttg	tcttaatgaa	tagaggagaa	tgactagtca
56221	ttattttcata	gatcataggt	acataggtga	ttttaaagg	tgagttagct	gtcccttgag
56281	ggttagttcaa	tgctcctctg	taacttgaat	tttttcatca	aattctttta	aatccagggt
56341	gggtcagct	ccctctgaag	tatcacaaaa	ccctaaatga	attagggtta	taagactaat
56401	tatattttcca	tagcaatggt	taagtggcct	gctgcccatt	caccttccaa	gccctgctct
56461	gttgggtttc	ttttttaccca	gtcatgatca	cgtaggacca	ggctgttttt	cgtagggggg
56521	tatggagtgc	caatgcctgg	ggtgttagga	atataggctg	ggctgtatc	atagtaccca
56581	cgccgctgcc	taactcttga	tatccagtga	ctgagtggga	ctttgtagct	gctttctctt
56641	tattaggaga	aaagagcttg	tttaataaac	ctaagaatta	atagcctgtg	ttcagtagtt
56701	ggattttgtaa	cctgaatggt	tttatgtcta	ctgacttgca	acgttgtcat	ataaattaaa
56761	gatcatagat	ccagctatgg	tttaaagggt	acttccagga	catggatttg	aaagatcaga
56821	atttgaatct	catctctttc	cattctagct	tggtaacctt	gaggaaacca	cttaacttct
56881	ctaaggctca	aagggtgtcaa	gtagattaat	gcaactatgt	atagctgacc	ctgaggaatt
56941	tttgggtccg	ctagaataaaa	gcttgaattt	ttggacatta	gaaggatctt	tgaagtatgt
57001	gtagcttctg	tgtgttaactg	aagaccagag	aaatttaagt	gagtttctctg	gaatagttta
57061	tgctctctgct	tctgtgtgca	tgattaccac	ccagaacacg	tgcttaattg	gggacaggta
57121	atccattcag	ggtgaagaac	atggcttttg	ggtcagctctg	cagtggatct	aaatcagccc
57181	cctgctgttt	actgtgtctg	tagtctagaa	aaagctattt	aacttcttgc	agattttagcc
57241	taactatgaa	attaaattgg	ggatttataa	aaactcatag	ggacgtggtg	aggcttgact
57301	gagataaatt	tatgcaaagc	tcttagcaca	atggctagta	cccagaaagt	gctcaacatt
57361	attattatttt	cctaattgaa	gggcattgat	gatttataat	aaaatggagg	ctgggtacag
57421	gctcatgtct	gtaattcccta	cactttggga	ggccgagttg	ggtggaacac	ttgagcccag
57481	gactttgaga	ctagattggg	caacctggta	aaactccatc	tctataaaaa	atacaciaat
57541	tagctgactg	cggtgggtgc	tgtagcccca	gctactgaag	aggctgaggt	gggagtagg
57601	attgagcccc	ggagggtcagt	gctgcagtga	accatgattg	tgccactgtg	ctccagcctg
57661	ggcaacagag	caagacccgg	tcacacaaata	catacatgca	tacatacatg	catgcataca
57721	tacatacata	catacataca	tgtaaataat	taaataaactc	atttcttgtc	agataaatgg
57781	ctgtatcttt	ataataagat	atctgtatcc	tgtgacttca	tctgtataat	aattttgtgc
57841	ctcttgtgct	ttctatgatc	tagggaggag	aaagctaatt	cttcttctatt	ttatgcacgg
57901	agcagagaca	cggagagcct	ctaattttatt	tcttcttggg	gtggccctgt	tttctgagca
57961	tgggtgtgtc	tgatccctgg	ggagagcaga	gccacactgt	ggatctgagg	tgctgggaag
58021	ccatccagtt	tctcctcctc	gaccctgact	caagtcttcc	ctgaaatctc	tgtcagcccc

58081	attctcttttc	tgtcagcccc	attctcttttc	tgtcagccgc	cctccttaca	taacccaaat
58141	gggttggtttg	caaagctagt	gtcgtctgagg	tggtctgtgt	acaacagaaa	agaatctagg
58201	gaggattcct	atgtgtcact	aaagccagta	attaagtgga	caacaggggg	agctaacact
58261	gaatgcacca	ataaatttca	agactcctgc	taccctagggt	aaccgagtc	atagtattag
58321	aaaccattttt	aagcattgggt	agtttttaaac	atgctgttta	aaaacaattt	taaattttacc
58381	ttcactttttc	taattggata	ctttactatt	cagagtacta	tgagatgagg	gtctcgcttt
58441	ggagtgcagg	ggactgatct	tggcttactg	cagcctcaat	ctcctgggct	caagcgatcc
58501	tcccacctca	gcctcctgag	tagctgagac	tataggctcg	tgccaccata	gctggcta
58561	ttttaattttt	tttgtagaga	tggggtctca	ctatgttgcc	caggctgggt	ttgaactcct
58621	ggcctcaagc	aatcctcctg	ccttggcctt	ctaaaatggt	ggcattatag	gcgtgagcca
58681	ccatgcccgg	ccaaacttct	tttgaaatta	gcttgttgat	tcttcctcac	ctcccagttg
58741	tttttgtgcc	agaattaatt	tttctccttg	tattatagaa	tagtttggag	gagtattgaa
58801	agaattaggg	ggtagagtgt	ccagatgtag	caaataaaaa	tacaagacac	ccagttaaat
58861	gtgaatttca	tataaataac	tacttttttt	ttgctataca	tatgtcccat	gcaatattttg
58921	agacatactt	atactaaaag	attgttaatt	atctgaaatt	cagatctaac	tgggcatcct
58981	gtattttatc	tggcaaccct	aactggggtg	gggatgggtg	ggagggcctt	ggatgtggcc
59041	agaggagagg	tgtcagagcc	tcaggtgtct	tcttgtgggt	gaccgaggtg	gctgcagcac
59101	agtagccttg	ctccctgggt	ctggggcctg	tgccttctcc	ctgtagtcc	ttagaatagg
59161	atgatggggg	tgggtcaagg	cagtgggaata	aattaatctt	gaaggaaacac	actgggccag
59221	agtcctagaa	cagtttactt	aatgatagtg	ttattttaat	tttcaattgt	ttgcctttct
59281	tcctctgtgat	acggaataaa	catgaaattg	tatctggagc	ggagcaggca	gaacttcat
59341	cttggttggtc	ttgttctgcc	ctcagcctct	gatgttctac	aaggtttatt	cttttgtttc
59401	tttttttatt	atccaaggaa	tgagatatgc	cagaaaaatg	tgacaggtat	ttatgaaatg
59461	ctttgaacta	ggtgagctta	gagcataagt	aatttttaggt	catttatctc	atcacacac
59521	tatctacaga	gttttaaccc	ttatcataag	gaacagacca	tgatgacact	gacattatca
59581	acataacgac	acacatgctt	ttctattcct	aatgctttgg	tgagagaaac	tggggcatcg
59641	gagaatgttc	tcagccatat	ttttgatatg	gcctaaggta	taatgaacaa	aagcttagat
59701	gagaaaagtc	catctgattg	atgcctggct	aattgacagc	caattatgtc	atgggtgccag
59761	ttcttaaaga	aattgaccca	ttaatccctt	tgtgtggaga	ggccagccag	caggcatctg
59821	ctttcttagc	atgcagcagt	actgaaaagt	ttattgaaat	aatcgtcacc	tgtcctttcc
59881	aaattctaat	tctcttgagt	ttaaactagt	tttctctaag	gaaagtccga	ttgactgaa
59941	atcacacatg	tctggaatta	tctctgagtc	ctttataaac	agaccaagac	ttggaagggc
60001	acaccttagg	ttacagagtg	ttttcctggg	ctgggctcca	aagcttccctg	actattgaac
60061	aataatgtgt	tctttccatg	ctactttata	tatttatatt	acagaagcct	tgtgcctttt
60121	atccgttcat	atccactttg	aaaacttact	tgtggcctgc	cagagcagag	gtacaaaaaa
60181	ttccggtatta	ttttctggtt	agaggcccca	cacaatacac	tatatgtgct	gaagatgagg
60241	gcactcctcc	tcctcctcct	cctcctcctc	cttctctatc	tacccttcc	tcacctcat
60301	ttttcttctt	ctttttttaa	aataatcata	agcatgtgtt	ttcttctcat	gtgcttgtaa
60361	atattttttgt	tggggattct	tgactgggat	ttcagaatat	cctgataggg	agaagttggg
60421	tatttcttca	tgagataaat	tccttaaggga	gagatttgaa	tagttttgac	ctttgggttt
60481	tctttctctc	ttcagttttc	ttctataaaa	atgttacatt	tcttgggtat	gagataaaaa
60541	caaacctata	atttgtgata	atgggtgaaa	atgtgattag	aattcacatt	ctaggtttta
60601	taatgacaga	ctacttatga	aagataagat	gtcagagctg	gaaggcttct	tagatattgt
60661	cgggttcaat	attcttttct	cattagagga	aatggagact	cagagacatg	aagtgacttc
60721	tccaagggtca	cacagtaagt	gagtgtgaa	gctgggagta	ggacctcttt	tgcctgactc
60781	caaacacagc	tttccccaac	tattgaggaa	aaggactcag	gacaatttaa	catttcaagt
60841	cattgaaata	tccttttaaat	gctcaaaact	taatttttaac	cttatgtgtg	tgtgtgtgtg
60901	tgtgtgtgtg	tgtgtataag	tcgttctaaa	gtacttaacc	ttctgaaatc	ttattttgac
60961	catgtagaac	acagttcgac	cttttttcaa	tctcatcatt	atcaacactg	ttttgtgaac
61021	atgggtcatt	gtggttttaa	ttcatgggtg	cccttgggct	attctgagtc	tataggactt
61081	gcccttagtt	acattaacac	tcccacatga	caaaactcat	gagtgcattg	ggaaactttt
61141	gatatacttc	ctcattgtca	gttgtcactt	tttgcactt	tcagaggtga	tatttatagc
61201	atttcccccc	ttctgtgttg	ttcctaccca	caggatttta	acttacagag	atgactgaat
61261	gatgacacag	aggggacaag	tccattgaaa	taagtcttgt	tttgttttgt	tttgttttgt
61321	tttttgggac	aggggtctac	tctgtctccc	atgctggagt	acagtggcac	aatctcagct
61381	cactgtagcc	tttgtctccc	aggttcaagc	tattctccc	cctcagcctc	ctgagtgtt
61441	ggacttacag	gcgtgggcca	ctacgctcgg	ccaatttttg	tatttttagt	agagatgggg
61501	tttcaccatg	ttggccaggc	tgggtctcgaa	ctcctgacct	caagtgatcc	tcctgcctcg
61561	gctgggtgtg	taggattaca	ggtgtgagcc	acagcacctg	gcagaaagaa	attttttttt
61621	attactcaca	tttcttaaga	gaagagggca	ttccatgcc	cacagggcca	ggaggagaag
61681	cacctatttg	ggtgaagagg	aagagatggg	agtcaggggg	aagccgaggc	cagagccttt
61741	actgggtttt	tatggaaaag	gcaaggcaga	ctggagggaat	cagcttgggg	ttggctaggg
61801	gctgggatgg	tcttttagttg	tcagttcctg	gccctgagag	atthagggca	ggggaaatgt
61861	gggctgggta	actgagagtt	agataaggag	gtggctcagc	tagatcacag	aggagatgga

61921	aactacttgg	ctgttaactt	gccctgtaat	tgatggatac	caaatagcca	aatacagatt
61981	cggagaaaat	gcaggacaa	ttcccaggct	tactttgctg	ccattcattt	tgtggtgaac
62041	caaaaaacca	cttacacaaa	cctgttgga	ggggtctctt	ttgtgagtac	tgtcaaagag
62101	agctatctag	aagagacctt	tttttgtgag	tggtgaggtt	taggctgaat	ttttattttg
62161	ttttggtgct	tttctgtaat	ttgggattat	taaaaacaaa	gactagacct	ttttatagat
62221	agaaacaagg	ctcttttatt	tggaaatacc	atgtgcataa	atgatgaata	aatagagtca
62281	gtgaggacct	tcttgccctc	atgattcatt	gttcttctcc	ttctcctttt	tctccttctc
62341	cttctccttc	tctttctccc	cttctcctc	ttccttcttc	ttcttttttt	ttttttaata
62401	tgagacacag	tctcgctctg	ttaaacactt	gtattagt	cattgaaagt	gtaatatata
62461	atctgaggat	ctcagacatc	ttaggaagat	gactgtcatt	tattatctat	taaccagggtg
62521	agcaacttcc	tggcgagtga	ggggtgcggc	agggaaagggt	ggagacgcaa	aggcacagtt
62581	ccctttgtaa	tgagagagctt	cagctcctgg	gaaggatctg	caatgcttag	tggggctgtt
62641	catctcta	atagttaata	attactgatt	tgtatgaagc	agaactgagg	gcagggaggg
62701	atgtgggaag	gccacaggag	atatgagttt	gcagccagag	tttaccggtg	atgaatcacc
62761	ttctgatcaa	agcaggaggc	tgggacctgt	gggggatgca	cttcgacctt	gggggtgtgac
62821	ccaggaagtc	agtgtttgt	gaagggcatc	ataggtcaca	gatgcctttc	caaagttacc
62881	agaacttggg	atccaattta	tttctcttag	tccaggctca	ggtttgattc	ccaatccctt
62941	gcattcacac	gagtcacagg	gcagaaaagt	gcagacgtgt	tcttgtgcga	cctccagatg
63001	tggtgtggag	atggaagatg	gtactctcat	cattagagct	gatttgcctt	ggaattaaag
63061	agagacatac	gcactgtgtt	tctgtggccc	atcatagt	ccactgatgg	tgccatgtgt
63121	catttggggg	aacattgact	tgtatttcca	ctcagtgtga	ggaaaggacc	ataagacaga
63181	attggagtaa	tttctggaaa	aaagaagtaa	atgcttagta	gagtgctata	cagtccttta
63241	ataacaagta	tttgataaac	atgattttgt	tatccatcct	tctgcaggaa	aagaagccaa
63301	gttaattttc	ctgagtttac	agattggagg	gttttttagta	taacctgtgc	ctttttcctt
63361	caccctgttt	cctcttttac	tactacagta	aagagggtga	aatttagttg	caaaaggata
63421	ccattgaaat	ttagttactt	ttgctcgtct	cttgctaaaa	gagttataaa	tgtgcagtc
63481	ttacttggga	ctgattttgg	tataatgtag	tgggtttcta	aaaatagatt	tctttttcat
63541	gtaattgaca	attaactcca	taagttactt	tacagaaatt	taagtttctc	tagaaattac
63601	tgcagtacac	attgcatgca	ttctccttaa	agaaaattga	cagaacaaaa	tttcatcctc
63661	tgtaggagc	tgtcttttcc	ctcacttgct	catctcatga	ggggaagcat	gtattatatc
63721	atgtaattga	cctcccagc	tgtatggcac	ccttgagtga	accaggtaca	accaggtccc
63781	acccaggcat	tttcttaatg	tccacaaagg	ctcgtgact	tcaagttagc	tatttgtgtc
63841	ctttaacttg	ttgccacaat	ttaaaaccag	gtgagcattt	tctgcacaga	gtggtcataa
63901	gcagtgtctg	ttctgtctgt	ctcggcctct	ttgtcacctg	ttccatat	gggcatgaag
63961	cactaggccc	atatgccttc	accattttgc	aatgttgttc	tgggacagag	ttatagggtt
64021	tttgccttga	acaaagcatc	tacattcttc	attcttaggg	agtgcagttt	ccattgcca
64081	tatgtggata	tcagttttct	tcaagcttgt	gtaccactcg	tatccactgc	tgttcagttg
64141	cataatctct	aagattaaaa	actacatttt	ggtaatgctg	gcaacgaggg	cacaaggaaa
64201	taaattgtct	gtttttataa	acatgtagct	actgatattt	ttttttgaag	gtgaaagcct
64261	tattttaaag	aggtctcat	agagagtttt	aaaattttag	aatgaaaaaa	ggctcttaaa
64321	ttattaacca	aatagtaaac	taattcattc	aaccaaaaga	cttactgaac	actccctag
64381	tgaagggtgtg	tgtgttagga	aatgtgctta	ctgggtgaac	acaacagagg	tgttccccac
64441	ctcatgatac	ccacagtcta	gggagtga	ggcaataatt	aagtaatttt	taaaaattta
64501	tagctgtaaa	gtaaagcaaa	ctatggcacg	tccatgaaag	agaggtaga	gcctatgagc
64561	ggggaggcgg	ggaccagatg	gagcctgggc	tcaggggctg	ccagacacac	acagctgcct
64621	acgggcagga	agggggccca	ttgggaagcgt	taaaacaagt	tctgtataga	aggaagcgt
64681	gtgaataaag	aaaaaaagat	tgacctttgt	gagttgtgac	atccaagggtg	tccaggaact
64741	agtttctcca	ctttcttttt	ctcctttttg	tcatggccgg	cagttgggcc	cttgtactgt
64801	tcatcagctt	caccatcaaa	atcaaatgaa	aagaaagaga	agggaaatgga	atgctcactt
64861	tctagacctt	actttaaatt	ctgttgaa	tttattaata	tttgggagag	agtttgaaat
64921	gataatccaa	aagatgtctc	cctttgaaca	tatgtagaag	ttaatcattt	agattaactt
64981	gcattttaa	acataattgt	atgtgatcgt	atacattttt	tggctcactg	ttttgtgttc
65041	aaaggcagat	ttcctagggt	agtgtcgatt	tattctactt	cttttaggct	gtgtaattca
65101	ccttttttct	aatttgggga	aaattgtggt	cactatagtt	atataatttt	tatttataaa
65161	cttgtaatat	ttatttaagt	tatctaccca	aatcttctc	ctgaaagata	atatcttttt
65221	ttcctttttt	ttttttta	ttcagtaggt	ttttggggaa	taggtgtcct	catagcttag
65281	ctcccactta	ggagtga	tataggatgt	ttggtttttc	attcctgagt	tacttcactt
65341	agaacaatgg	tctccagttc	catctagggt	gctgtgaatg	ccattatttc	attcctttat
65401	atagctgagt	agtattccat	tatctatctc	tgtatatgta	tatatcacia	cttctttatc
65461	catctgttga	ctgataggca	ttgtgtctg	ttccatattt	ttgcagtgtg	gaattgtgca
65521	gctataaaca	tgYgataaca	cagtatcttt	tcagatatgt	gtccgtctga	catgttctgc
65581	accaccact	acactgcagg	gcatttaggt	aagcagcatt	tataagagtg	actggaaatg
65641	gcttaggaaa	aatggtacag	aatgtaattt	atcagaataa	tcactgatcc	tcttagcatg
65701	tgtattcaac	aggtaactgt	gactctaaaa	tgttatttgg	gagggaaatt	gcaaccaaa

65761	ttagaaacag	ctctagaggt	gctgggatct	aggggcctta	gattttttca	tctcagttaa
65821	tcacaccatg	ctttaggaag	ctttaggata	agattctggt	gctagtgtcc	ccggttggtt
65881	ttgcttctga	gaaccaattt	tgcagttgtc	acatgatgag	ttggatcctc	atgcagatcc
65941	tccaagggtc	ttactgaaca	ccatcggttt	atgctagtga	gggtttgtgg	ttctcagctt
66001	cccgcattgg	aatgaacatt	atttgggcct	gagctgtgac	agcttccttc	ttcagcttcc
66061	ttccttggtt	tcttctggtt	gggttatgga	ccacttcttc	tggagaacag	cagggtggtg
66121	ttcaagcacc	atcgctactc	cctagtctag	agctcagcca	cccagtacta	tagccaatag
66181	agatatattg	ctcaaaaaca	catgagagat	ggtagtgtga	cagaataatg	gaattgtaaa
66241	ttgtatttag	ttttaatcaa	aactttaaaa	ttgatatcca	atataatga	ttggaacaac
66301	ttgaatatgt	gagcttacct	tttcatctct	aaagtttatg	ataaaaagatt	tccagtaaaa
66361	atttaattgtc	caagttgaga	tatgctgtga	gtgtaaaata	catgctatat	ttcaaagact
66421	tagtatattaa	aaagaataat	atacaataag	taattttaaa	atgttgatta	tttgttgaga
66481	tgataatatt	tttggtagat	taggttaaat	aaaatgtact	attaaaataa	tgttatcttc
66541	ttaaaattgt	gatgactata	aaatttgaaa	ttgcataggt	ggcttacata	atgtttctgt
66601	tattgggcat	tgccctagag	ctaactgagg	aaaagatcat	agggcaccat	ttgccattgt
66661	tgtagtatgt	tggctttcag	tactaggagg	taaagtagat	actctttcca	tactaaatac
66721	taaatgccac	tatcattaaa	aaacaacatg	gcaaatcttg	cccttaaggg	gacctcattt
66781	tttccccagc	aaaacaaaac	cgacagcctg	ttatatagct	aaagagtaga	tgaaaatact
66841	taaaccaatat	aaatgaactt	tattggttaa	taggtgcagc	aaaccaccat	ggcacacatt
66901	tacctatgtg	acaaacctgc	acatcctgca	catgtatccc	agaacttaaa	atataaatata
66961	tatacaaaat	aaactttatt	agttatgcaa	tgaaataaaa	cacagcaagc	aataaccttta
67021	ggttcactga	atttttaaagt	taacaccttc	aatacaaatg	tatatatttt	gtatacacgg
67081	ttatgaaggt	atgaacatat	taatgagcaa	aataattatg	aatacttgtt	cctcttttgt
67141	agtttttaaat	atgtataaagc	aagaagtaac	ttgacatgac	ataatgctgt	gccaccttgc
67201	ctattctgtc	gtgggctcca	aatgtaaatt	catcagaaga	gctcacagct	ttgttgaatc
67261	tctctcctct	ttggggagat	aggctcatat	gtctcaattt	tgaagcacc	ctcttggaa
67321	gaggtcagtt	agaccactcc	aaggaccact	ctcttttctt	ctttgtattt	tctggggaaa
67381	acgtatagct	agcatgctgg	gactgtttgt	tttgaatgag	tttgtttatg	agttttcaaa
67441	acaaaattat	gaaaaaaagg	gaagtttcca	actccttgtg	catgtcttgc	tggctaagca
67501	gatgatctct	aaaaacagat	tacaatatca	ttttgaaggt	aacaaaggta	tcttcttgc
67561	ttatctttaa	ggttgcatat	tttatgtgtt	atacattggc	caattaggaa	ctcttcttta
67621	agaaagacag	tcatttttaac	ctattgaagt	cacagaatga	aatgatgaag	taatcgtagt
67681	tgttcccagt	ttagcatata	tgtatatata	tgcattgatg	aataatatga	tttgttgtaa
67741	acaaatgaaa	aactgcagaa	acctgtaatt	tgtacattat	tatttcagat	caccgtaaca
67801	aatattacca	ggtttttaat	ttctttttaa	aaaatgcatt	tctagggctg	ggcgtgggtg
67861	ctcacgcctg	taatcccagc	actttgggag	gccaaaggcag	gaggtacacg	aggtcaggag
67921	atcgagacca	tcttgcccaa	catggtgaaa	ccctgtctct	actaaaaaga	caaaaattag
67981	ccgagcgtgg	tggcacgtgc	ctgtagtccc	agctgctcgg	gaggctgagg	caggagaatt
68041	gcttgaacca	gggagtcaga	ggttgcaatg	agccaagatt	gcggtgcac	tctgcctgg
68101	tgacagagcg	agaccccgctc	tgaaaaaaaa	aaaagaaaaa	aagattttct	aaaattgtat
68161	ttatactctc	tgtctcttcc	ccatcagcca	tcaacgcttc	cctccctccc	tctcccttcc
68221	ttaatgataa	gccctcagct	ttgctcagga	ctcagcctcc	catgtggttt	tggtaagtgg
68281	tctaagacct	gaggcccaaa	gcgtgattgg	ctgatgctgt	gatttctcag	cctgggttgcg
68341	cattagaatc	accctgggag	ctttagaatc	cagatgcctg	ggctctactc	acagagattc
68401	tgattttaatt	ggtgtgttgc	agaacctgac	ttgagccatt	tgcgaatggg	tcattgatag
68461	cttgtagctg	taacgattca	aacatataca	acataagcag	ggtgaccaca	gaagttattc
68521	tctaagtctg	gtatgttctc	aaatgtcctc	taaattctta	tcttccactc	cagggctttt
68581	gaagtggcct	gatccaaaca	ccttctttcc	gacattaaaa	acattagccg	gttattttgc
68641	ctcatcagca	cttctctacac	ttccttaggt	gagccaggtt	gcttttaatc	tcttggagct
68701	gccttcggtt	ataggccctt	ttttttctt	ttttgtttta	aattgtatgt	attgaaagta
68761	tacaacatga	tatttgatat	acataattct	agtgaagtaa	ttactacaat	taaatataca
68821	cacccatcat	ttcacatagt	tacctttctt	ttttgtggc	gagagtacct	aaaatctact
68881	atcatagcaa	atttgtaata	tataatgcaa	tattattaac	gacaatactt	gtgtggtagt
68941	ctagattttt	tcattccaaga	taactgcac	tttgtatgct	ttgatctata	tctccccatt
69001	ttccctgccc	ctcccgat	tgccctttga	ggcagtatag	tgtgacagtt	agccatcact
69061	gaattctggg	ggccttttgt	ctattccagg	gagatgaaat	tgttagagaa	agattcaggag
69121	aaagaatatg	aaaaggactt	agagaggata	caaaatcatg	aagtcacaaa	tacctgccaa
69181	acaaatccac	ataaggagaa	aataaagaaa	tgtcagattc	ataaaaagatc	aaaaaaccca
69241	aactgttctt	agagggaaaa	gcattgcacag	ttagggaaac	tttttttaaa	aagttttaat
69301	aaaaatctgat	ctctagtggg	aaatactggt	tcccaattac	acctagactt	ttatcgtgaa
69361	tcaggttttc	tagttgacat	tgttgttttt	catttgaaaa	tgaatgctgt	gtattctcct
69421	tgtttcaatt	cccttatgta	tttttgtgtt	aactccctct	gccggggagc	gccagactca
69481	gatgaaaggt	attatgacat	tcgagatgaa	taatgacgta	agaaggactt	actctattgt
69541	atatcacggt	gcagtatcaa	aacattttgt	cccatgagga	gaggcagtga	gtcagaaaa

69601	cctgtttgaa	ttgtattata	atgtagaaaa	ccatttcaga	attactgtct	gacatttggg
69661	cagctgggac	ttttagactca	ttccacggcc	acccacacct	agacatttta	ttaggaagat
69721	gctattcttt	tttagggcta	tcactggatc	tttgcctca	ctgatgaatt	aattaggtag
69781	gtaataatga	atctgagtta	ttactgctga	caatttagtc	ttattcttaa	aaacgttcat
69841	ttcatgggtg	gttactctgt	tctgtgttgc	aataaaggaa	tacctgagac	tgggtaattt
69901	ataaagaaaa	gagggttttat	tggctcatgg	ttctgcaggc	tctgcaagca	tggcaccagc
69961	atctgctccg	ctacctgtga	ggcctaggaa	ccttgaatc	aaggcagaag	gctaaggggg
70021	agcaggcaag	acacacggcg	aaagaggag	caagagagag	caaggaggga	ggggccttgc
70081	tcttttaacc	aaccagctct	tgtgtgaact	cagaatagga	actcacttgc	tatggcaagg
70141	acaggaccaa	gccattcatg	agtgatccgt	ccctgtgacc	caaacgcctc	ccactaggcc
70201	tcacctccaa	cattggaggt	cacatttcaa	catgagattt	ggagggggaa	aaaacctcca
70261	aacctcatca	catgggttaag	tggcataaga	tgcaactcta	ggaatttttg	cttttttaaa
70321	accagcttc	gagctccctt	ccattccctc	atcaagatat	tcatgatgga	tctgcattgg
70381	gacctgcttt	tgtcttattg	tttcccaaga	ctaaaaaggg	ttttcagcgc	aagtgttctc
70441	gtaatgtttt	tgtccccaag	attggatggt	tttcagttgt	attatctccc	catatttctg
70501	acatgtacag	acgtgcccag	ctatgataKt	ctgtgtggat	ggtcttaaaa	atgagttatc
70561	attatgaata	gtttgatgga	aagatttccc	aggaagcatt	atgctttgag	tgttggggaa
70621	gagaagccac	cagatgtcgt	gtaggccctt	gtccctttac	tgcctttcct	ctctgtttgt
70681	accctgtac	aaggcagctc	ttgtggttac	gatttgttaa	caatttgagt	agctccctta
70741	tattttgctt	ctagagtaat	taaacattgt	ttcttagaaa	tgtaaacaatg	ggacttagaa
70801	ttacgatgca	gtaatgcatt	tggcagctga	actgtgtact	gggaaaagaa	aaatatcctg
70861	gagcgggttt	aagtttcgca	gataagaaaa	atggctttat	ttagtgaatt	ggaacaatac
70921	aagacctgtg	ccctgaagtc	agtactttct	cttcccagat	gagtttcccc	aggcacaaga
70981	ccaacttcta	aaataaaaact	tgaagtaaa	tggcaaaaac	cagaaacaga	gccctgtaaa
71041	ctagactttt	atgctacggc	gccataaaaa	taacttattt	gctatcaaaa	taagcttaaa
71101	cagaaccttt	tcttgacaga	tgacttttcc	tctcccctta	acctgacagc	acccccaacc
71161	cagtcttttg	tcataagcct	gacagcacgg	tgagatacaa	gttcctagtg	gcagtggagg
71221	ctgatgttta	ttatttagag	agtttctgaa	aatgaaagtg	taatgtcttg	agtcacttac
71281	ttccaaattc	ttcccagaga	actttaaata	gtgttgtatt	aaaaaacagg	agtaaacat
71341	cattcgttaa	acacacccca	gaataaatct	cttatttctg	catgaaggca	actgatctga
71401	aacatttttt	ctatgtgcct	tttagagacg	gcaccaaat	tcatgcgctg	tcgtgtgaag
71461	agaccaccaa	acaggctttg	tgtgagcaac	atggctgttt	atttcacctg	ggtgctggtg
71521	ggctgagtc	gaaaagagag	tcagcaaagg	gtggtggatt	atcattagtt	cttataggtt
71581	ttgggatagg	cggtgaagtt	aagagcaatg	ttttgcgggc	aggagtggat	ctcacaagt
71641	acatttctaa	gggtgaggag	aattacaaag	aaccttctta	aggggtgggg	agattacaaa
71701	gtacattgat	cagttagggt	ggggcaggaa	caaatcacaa	tgggtggaatg	tcattcagtt
71761	aggctatttt	tacttctttt	gtggatcttc	agttacttca	ggccatctgg	atgtatacgt
71821	gcaagtcaca	ggggatgcga	tggcttggct	tgggctcaga	ggcctgacaa	taacacagt
71881	ctcaaaactg	tcagatagcc	tgtgtcagg	ctgaagatgt	gattttggtt	gtttatactt
71941	ggagtctttg	gatgggaata	gccctggagt	cagcccttca	tttcaggcag	aggagcagag
72001	gagtgtgagt	tagacacaac	tttggatcag	gggccaggaa	aacttggttg	tcattcctga
72061	ccagttacta	actcgtctgt	tgggcaagtc	acttcatttg	tctaagtctc	tctttttcct
72121	tttgtaaaat	aaggatgtcg	gcaaatctgt	gtggtccctt	ccaacagtgt	tttttaagtt
72181	ggtgcctgag	tatctgaagc	aggagatacg	aagggtcatg	tgagctgcac	atccttattt
72241	gctccgcagg	gaggctggct	aagacaccgc	gccctgcct	tgtaaccatg	aacaaggttc
72301	tcgcaggggc	tcgtctcaac	tgcagggtcc	ctgaagggtc	catccttttc	tttgctagag
72361	ggaatttgga	tgtcgttgg	ccttgccata	cccttgtctt	tgaagatac	agatccaatc
72421	tctgtgtagc	agtttaagtga	tctgactcag	acataattac	tcagtcttct	tagagaatga
72481	gaaaactctt	ctcagaattt	ttaagaatgt	tcttgaagga	caataaaaagc	tctcattcag
72541	gataggcccc	aaaacatttt	tttctttata	atgtggtgcc	atttcctcat	tttgcttttg
72601	ttcatttggt	tattccttca	acaaatatatt	ttgagaattt	gctgagcact	aggtattact
72661	agatactagg	acagtgaatg	aagtaagata	cagcccttat	cttcaataag	ctgtatgcc
72721	tgataatgat	acccttagtg	tcttctacaa	gctatacgg	catgcatcac	ttaacgacag
72781	ggacacattc	tgagaaatgc	atccttaggc	cattgcattg	ttgtgtgaac	atcatcgagt
72841	gacttacaga	aacctagatg	gcgtggccta	ctgcacacat	agggatatg	gtgtaacctt
72901	ttgctcctag	tctacaaacc	cgcattggcat	gttaccatac	tgcatactgt	aggcaatttt
72961	cctctccccc	taacctgaca	gcaccattta	cctgacaaaa	atggtaagta	tttgtgtatc
73021	taagcatatc	tagacataga	aaaggtacag	taaaaactaag	gcataaaaaag	tgaaaatggt
73081	acacctttac	agggcagatc	catttacgca	accaccactg	catgctggt	ccattgttaa
73141	tggaaactgt	tgtgtggggc	ataactgtat	ataaaagtat	agctacttta	attttaggtt
73201	aaccttggt	tgtgggaaaa	tttgtcttct	gtgctgtctt	gcactgaatt	ttgcattggt
73261	atttttccct	taatagtggc	tgcaaaaaaa	cttataaata	cagaaccttc	ttcattttata
73321	gaattcttct	gcattgacct	ggaaaatgac	gttgagaatt	ggacgttaca	ctacaatgta
73381	gtcctccaat	gaaggctctc	aatgggcatt	tctttaaggc	ctaagttaaa	gataaaaatg

73441	aacaacttcc	atcactacaa	aagatagtg	actcggagga	acttgtagag	atTTTTTTTT
73501	tctttagact	gttttttctca	ctactcagg	ttcctttttg	agttttgccc	ctggaggctc
73561	agagttgaat	tctgttggta	gtcacttaga	acctttctac	tgctctgtct	ttcctcagtt
73621	gtgttttccc	catgtgggtt	tgTTTTGGGA	aagcagtgga	ggggaattcc	tcttaggttg
73681	aaataacttt	tagagcgatg	gtgccacagt	ttacaaatat	ttttagaaaa	atcctgtcag
73741	attcttggga	acttcagact	aacttcacat	ctaaaagtct	cttttctttt	cttttctttt
73801	ttccttccct	ccttccttcc	ttccttccct	ccttccttcc	ttccttccct	ttttttgggt
73861	tttgagacag	agtcttggct	tgtaaccag	gctggagtgc	agtggcacga	tctcagctca
73921	ctgcaacctc	tgccctccag	gttcaagtga	ttctcatgct	tcagcctccc	tgagtagctg
73981	ggactacagg	taccctccac	tacgcctggc	taatttttgt	atTTTTtagtc	gagatgggg
74041	ttcgccacat	tggccaggct	ggtctgggac	tcttggcctc	aagtgatccg	cctgttggc
74101	ctcccaaagt	gctgggacta	caggcgtgac	ccaccacgcc	tggcctaaag	ttcttattta
74161	aaaatttttc	ttctgatttg	ttagttaaag	aaggtaggtt	tgaagcagtg	accaggaatt
74221	ttcgggaaat	ccattaagga	ataaattatt	cagtaaaaca	gtctcaaagt	gagggccaga
74281	gtgcaggaca	gaggcagaga	gagatggtag	cagttataaa	agagaagata	cttgattaga
74341	gaaatcattg	tcagagtaac	cttatgctta	gaaagaaatc	acacgcgaag	ctctgtgttt
74401	gaaatcagaa	gggaagggtg	gcattcggat	gaagaggctg	tcggacttgc	attactttga
74461	ccactactgt	tgTTTTtGCT	gttgtgggtg	ttgttgttgt	ttggaagatg	gagtaaattg
74521	caagcctggt	ggatttcatg	tgTTtagaat	tgtagctaaa	atagctcctt	acgttgaagc
74581	atTTcctcaa	ttctataccc	acgttctcag	tccttgtgtt	actaccagta	actcatTTTC
74641	caaaatgcag	aatttgcattt	tacatttttag	ttctttcaat	atTTtgatca	aatacatgtt
74701	cagtggaaag	ggtatgttta	atTTcctttt	ggttcggcac	taatttaa	tatgaataga
74761	gataacgtaa	ttctagtatt	cgtatttgat	tgTTaaaata	tttgactca	gaagtgcatt
74821	tacacgtttc	caaatttgac	aagtaggaaa	aggtatagag	tccaaactcc	ttcccaggcc
74881	tatccacagc	taccacagtg	cctaatcccc	ggcaaccagt	gttaccagtt	tcttgcata
74941	ccttctgcat	actaccaata	caaaagcatg	ttattctgta	gccccagcta	cttgtgaggc
75001	tgaggtggta	ggatcactcg	agccgaggag	ctggaggtag	cagtcagctg	agattgtgcc
75061	actgaactcc	agcctagggtg	acagtaagaa	cctgtctcaa	aaaaaaaaaa	aagcatataa
75121	ttttgttctt	cttgttttta	tgctgatgag	gacatgctat	gtccactgtt	ccatggtag
75181	tgcttttata	ttcagttatt	cactaaaaaa	atgagaagta	tagagtaaaa	tgaaagtctg
75241	tactattctc	cactagaaac	aggtgtgggg	ttcaaggagg	agtggtgatt	aattattgca
75301	aaacttcttg	ttaggtcatt	aatcagaacg	ggagccatgc	agtggtgaag	ggccctggct
75361	atgatctgaa	acatttttct	gaagcagtg	tgagttctgt	aactgaagtc	catggggact
75421	tgctattggc	ataactatat	gtggaaaatt	ttgctttttg	agtaggacat	tcccattgct
75481	cacattta	ttgaattttg	gtatgactgt	actaagttaa	catagctttt	cgcatTTTAC
75541	agagacattg	agtaaatcat	tgactaaaga	taattaaactc	ctttatatca	Yatgtgaaa
75601	tgttatgggg	gacatatcat	tacctttctt	ggcttatgga	agagatgtaa	tctaattcat
75661	tgagtctgac	tgTgtTTaag	ctattgtttt	gggtgtcatc	gttggtctct	agaacaggga
75721	ctgggcctta	gtccttttcc	ttgctcatac	ctggcctcct	cctagaagac	cctgcttttc
75781	atgttttatg	cctcagaacc	aagatgtttg	gggtcccaaa	gtagggatgt	gtatgagcac
75841	atTTTTtgat	actttcatct	ttctcttact	ctctaacatg	cgttccgcta	gtgtcatgta
75901	aatacagtga	atcagatatt	tctctgcctg	caaaatgcct	tttgccagta	tgTgggaggt
75961	gtgttgacag	taccacagtg	gattttctgc	atctgttcat	acatccatct	ttcccactta
76021	gctgtgagtg	gcctcagtg	agaatgggtg	cttaactcta	gagtcgtgta	ggaaatgcca
76081	tgacacattg	tcaaatgaat	ggaaatgcag	ctgacttgct	agagctcagt	tctgagtcca
76141	gggtgcttga	atgctatata	atcaagttga	gatcttagtc	tggggcagca	gaaaaaaga
76201	aaagggttact	gagcaggaga	gtgacaggct	ttcatttgta	tcttaatggg	ttatttgag
76261	tggcattaga	aaaggggaga	gattggggta	aggcattagt	tggaaagctt	tgataaagtc
76321	taggtgagtc	aggaaggggt	tctgcattgc	ttgtttccca	ttgctctaac	tagatttctt
76381	agaaaaaaga	aaatcttaca	ccagactttt	acagcgtttg	caactgagta	aatttctcct
76441	tgtaataccc	tggatattat	acaattatat	aaagcgcagt	ataataagag	agttagtagc
76501	ttcattgtaa	tcagtaaaaa	taatttttaa	gacaattgtt	atTTTTtgta	tattgtatta
76561	taacaaaaaa	gaagtcacat	agtgggagga	gggtgtgttt	tgagacgttt	gctttcatct
76621	cttaatcctg	cagacattta	tctagtccct	tctctgtact	cttaaatgct	agggatttaa
76681	aagacggata	agatacaggc	ttattatgtc	tgtgttccag	acactggact	ccaaacataa
76741	agcaaacatt	acattattct	tccttatgta	atagaaatgt	ttatgttaaga	ttgtgtgtaa
76801	atcagtcttg	aataaaactga	attaaatgga	atgtgccaat	aggagttgct	atttagagaa
76861	gccctgtgat	aaaacatttg	tataataata	tacttatcta	aacaagccca	ctaatttcta
76921	tgggtttttt	ttttttttcc	ccccaaaggcg	aggtatcctg	tatattctgt	gttgattgaa
76981	ttccagttgg	ccaaccacct	gcctgggttaa	tagagtatac	cattaactta	gtgacacatg
77041	aacctgtcaa	aggaaataag	acacaatcta	tcgttgttca	gtctccacta	ttaaatagat
77101	tttcattcac	ttcagcttgg	gtgggtgtaa	tttgcatctt	tcctaacagg	caagcatctg
77161	caagtgtatt	gcctgtagaa	gctcattaaa	aatcagtcca	aatcctgaca	ctgtctctag
77221	aagcaggcat	taacttgcag	aaagtgggtt	gtatttccag	tgtcaKtaga	gcttctcacc

77281	tcttttctg	cttccatgca	agtttagtcc	taaactagta	ccactatttt	agtacaacta
77341	gtaccactca	aataatgctg	cttttttaaat	aaattcaagg	ggaactgcta	aggaactgag
77401	aacctgtaag	gtgacaggaa	aaaggaaatt	ctattttttg	gggctagttt	gtgtattgaa
77461	aataattttt	gctgagaatc	aagctaagaa	aattacttgc	taatttaaac	aacatgacag
77521	tcctcagaat	tttccagcaa	cagttaggag	caactgtgata	aagttggctt	ttctgttgag
77581	aacgttttac	ctttttgctt	cagcttcttt	aaagagtttg	aaattagtaa	tttcagtaga
77641	gcagctttgc	tggtgctgcg	tactgctcag	agcttagtga	gctgaagcct	tttgggaaaa
77701	tagcatttgg	ggagagactc	gtggtgtaat	agctcatccc	actggcacat	gtcccagagt
77761	aagctgggct	ggaagctttt	agtgtagtta	aaagatgcca	gtctgtcatt	tgcatgcact
77821	gtaattgggc	aagtggtttc	aggctgagct	ttacattatc	cttccactga	gagcagctgg
77881	tggtgggctg	tagattccat	atgagctggg	gacttatcat	ctgggtgtgt	tagtgcaatc
77941	ctgcctcatc	ttgggagcaa	ttttttattg	aatgataata	ataatgcaca	atcttgggtga
78001	aagataatgc	ttgtggttat	tagaatgtgt	agactgagta	ggggcttttg	catagagatg
78061	gtaatgggtg	ggaaagacat	atttaataaa	aggattgtaa	tggggagaaa	gtaaatattt
78121	tgacagataa	ggaaagccac	aaatatgatt	aatttaagag	tcttaaataa	aatgtctata
78181	aaatgttaga	tttttagtac	cagtaaggca	aagtggccaa	tctctagctt	cctttataaa
78241	gtctactcat	ccttggaggg	tcgctttttg	gtgctttttt	tggttgttca	gagtcaactt
78301	cttctcttag	ggatatgttt	aagtctgtca	cctttcctta	ggaattgtgc	caatctgatc
78361	atltgttcca	ccactgctgc	ccgccccctt	tttcttgact	caaagaataa	tttgtgtacc
78421	ctgtctctgt	tttttttttt	tttctttttt	tttttttttt	attatacttt	aagttctagg
78481	gtacatgtgt	acaacgtaca	ggtttgttac	atatgtatac	atgtgccatg	ttgggtgtgc
78541	gcacccatta	actcgtcatc	tgcattaggt	atgtctccta	atgctatccc	tccccctcc
78601	ccatctctgt	tttttattat	ttatttattt	atcgttagag	acagagtctc	actctgtcac
78661	ccaggctgga	gtacagtggt	gcagttatag	ctcactgcag	cctcaaactc	ctgggctcaa
78721	tcgatcctcc	cgccctagcc	tccctgagcag	ctagaagtgc	aggcacatgc	caccacaccc
78781	agacaatttt	tctatttttt	atagagatga	ggctctctcta	tgtttcccag	gctggctctg
78841	aactcctggc	cttaaatgat	cctcctgcct	cagcctccca	aagtgtctgg	attatagggtg
78901	tgagccacta	cacccaggcc	ctgtttttta	ccttgcaatac	cctttctgca	agattgaatt
78961	tatatattag	tataaaagtt	gtggacaaaa	tagaacactc	cattataaaa	gcctccttca
79021	tttggtttgt	tctctggtgt	ttgatttgac	tgatgtggat	ttgagtatgg	aagtgttcca
79081	tgcccttatta	aggaaagcac	tttgggaatt	ggccaggggc	cacttaactt	taagtctagaa
79141	tggcagcact	ttggcagtc	tcagtttctc	ttatttccta	acccatccct	actcattaag
79201	acggggctat	tgcatctccat	tttcagggaa	tgctctttca	tttttcgttg	gtgagagaac
79261	atgaatgcct	cttaaaatgg	tggtttgagc	ttgctgagaa	ttttaggggg	tccacagagt
79321	tgaaaagctt	tacaggctat	cagtgaagca	tgagggtttt	tcgttatgaa	aatgtcctga
79381	gatgggggga	agactggaca	gatgaggtag	gggagcctcc	ttgcaaaagt	agaattcagc
79441	tgttttatact	ggtaacagaa	tctgctttag	taaggatgaa	gcaaaaagaa	aaacgatatt
79501	aacaccttga	gaaaatctct	gtattgtgag	cttaatccaa	caactccaac	gatgttagct
79561	actttttcaa	aatacatctt	agccctgtgt	acaataacat	ttacttgtgg	agtgaatttt
79621	tgtattatgg	agattcactg	taatcagtaa	tcttccctcc	attgagattc	ttcctgtttac
79681	tttcttattt	aaaaaccttc	agtagcctgc	atcctcaggg	ttctgtttac	agagacccct
79741	aacagtctat	tcagccttac	ttctgccatt	atccctactc	taggctgagt	gttaaccata
79801	cctacctatg	tacatttgca	gctgtgccat	tttatctgtt	tggtattctta	acccctctc
79861	ttgcccaggt	gcctagtggc	gccctgtttg	cacgttgagg	ggtagcttaa	gtgtttactt
79921	tttcccagac	tccctcagtc	cgtgttagca	gtaccctctt	ccatactctt	ctattatctt
79981	ctgcatactt	ctgtactcca	cctatctgtt	gccatgaaac	aggtcaccct	aaaacagtg
80041	cttagaacia	tactatggac	tgctgtttgt	tccccccacc	ccaactcata	tggtgaaatc
80101	ctggccccca	gggcgatggt	attaggagat	ggggcctttg	ggaggtgatt	aggtcgtgag
80161	ggtggagccc	acatgaatgg	gatgtcttta	taaaagagac	cccagagagc	ttctttctct
80221	ttctaccata	tgaggacact	gtgagaaaac	atctgtgaag	cagaaaagtg	ggctcctacc
80281	agacacataa	tctgcttgct	ccttgagccc	ggactccttg	agcctggact	cctcagcctc
80341	cagaactatg	agaaataaat	gtgtgggKtt	tttttttttt	tttttttttg	gtcttgtttg
80401	tttggttgag	acagggctct	tctctgttgc	ctgggctgga	gcaaagtggc	atgatctcag
80461	gtaactgcaa	cctctgcctc	ctgggctcag	gtgattcccc	cacctcagcc	tcccaggtag
80521	ctgggaccac	gggcctgtgc	caccacgcgc	agcttatttt	tgtatttttt	gtagagagg
80581	gtctcccat	gttgcccagg	ttagtctcaa	actcctggaa	tcaagtgatc	tgcccacctc
80641	agcctcccaa	attgctggga	ttacaggcgt	gagccaccga	acccagccag	atgtttgttg
80701	tttaagccac	tcagactatg	gtatttttgt	atcttccatt	gtacttaatc	tatgggtctt
80761	atattcagca	gtattctatg	ttatgtaagg	agttgtggcg	tgaaccactt	tccttctagt
80821	ttttatgttt	ctttttaatt	aaatttttaag	cctgggaatc	ttggtaatga	catattatat
80881	gcaaaatatg	taatatattga	cttgattttc	atcagatcag	tttttagcact	tttcagtgtg
80941	caaaggtggg	tggtaatggc	tttttccaaa	gcatagtccc	ttgggtgtat	tacactatta
81001	atgactatgg	gtagtgtcag	accccgagcc	caagccaagc	catcacatcc	cctgtgactt
81061	aaacgtatac	atccagatgg	cctgaagtaa	ctgaagatcc	acaagagaag	taaaaatagc

81121	cttaactgat	gacattccac	cattgtgatt	tgtttctgcc	ccaccctaac	tgatcaatgt
81181	acttttgaat	ctccccacc	cttaagaagg	tactttgtaa	ttctccttac	ccttgagaat
81241	gtactttgtg	agatccactc	ctgcccgcac	aacattgctc	ttacttcac	cgctatccc
81301	aaaacctata	agaactaatg	ataatccacc	accctttgct	gactctcttt	tcggactcag
81361	cccacctgca	cccaggtgaa	ataaacagcc	atgttgctca	cacaaagcct	gtttgggtgt
81421	ctcctcacat	ggacgtgcat	gaaaggtagc	ataatacaat	ttttgctgaa	gtttgtgtgc
81481	tcttctactt	ttaaattaag	gaccagatag	gaaaggacat	atggtaatct	atcaaatata
81541	tagccattga	tattttcttt	gttttctttc	actaacgggt	ttatgcaaga	tttttttaaa
81601	aaaacagcaa	aatgaaaagg	ttcgtaatat	cctgacttcc	tgttttatta	taaatgagat
81661	gaacaccagg	tagtggttaat	ttcctttaat	ttctttatga	gtcatcctca	aattaatgtt
81721	tgaggaaatg	ttgggtgttt	tcaaatactc	ccgacatctt	ttagtttagac	agcatagtga
81781	aatggaaaga	aattttgaat	ctgctagact	caagtttgca	ttttggctct	gtcatttact
81841	catagatgta	ctacttgag	taatcatact	gattttgtag	accaattgga	aagataattg
81901	accagtgttt	tagggtatat	gtatatacac	aaaatttgac	agacatgtaa	tcacacaatg
81961	cttaagcccc	atatgttatt	aagcccacac	tttggtaaa	aacatttggg	gcttaagcat
82021	ctgtgtgatt	ctatatatgc	cggtatttgt	aaatgtaaaa	gtgataccag	ctaaatcacc
82081	catctttctg	ggaactaagt	acaaaggaa	ggttctgttt	ctaacccttt	ggcttagagt
82141	atcatagagt	tttaaatttg	gccagatttt	aatttgacag	ataaaaaaat	gaggcacaga
82201	taaaaaaatc	gagaagttga	atgaattatg	aaaactcact	agaggcagag	ttgaggacgt
82261	tgtccagtgc	tctgttgtat	cattatgtca	ctgcctgccc	tactgttttc	ctgctcttcc
82321	cctcttcttc	ttcctccctc	tcagccctgg	tcacactacc	ttagtgcagt	cacacacgta
82381	tggcaaagat	caacaaggat	gctgggttca	gggtcctttt	cccactggag	tctggcaaca
82441	ctgttcctgc	cttgatatgg	ttcttggttg	ttggYcactt	aaaacaatgg	catattttga
82501	agggttgaat	agacttcctt	ctttttaagc	ttttcttttt	atttatactt	tttgatccg
82561	tgtaatttct	ctgagtcac	cacattattt	tctacttttt	ctcagtttcc	atgacctgta
82621	ttaccattac	aggtgtctct	gctataattc	aataacata	ttcctgaaaa	cctcatgttc
82681	tacaaaatca	tacattcaga	ataatctggc	taatgggaaa	tattggtttg	gggcagtctg
82741	gcttatggga	aatattgggt	tggggcagtc	tggcttatgg	gaaatattaa	gttgggacca
82801	accactctaa	atctatgcta	ctttgcaagc	acagcactaa	caaaaaaat	gaaaaccatc
82861	ataaaacagg	agcacagttc	agaagacata	ctacattcct	actatataca	gatacacttt
82921	ggtaaatatg	actttaactc	atgaaaatat	gcggggctgc	ttgatggaa	gggatgttaag
82981	gaaggatatg	aggctgggtg	gtactgggag	acaagaatga	aacacaccaa	gacgtttgca
83041	tgagatcatg	caaagagaat	catgcagaag	gtacatctaa	gacacaaggc	caaccgggca
83101	tgggtggctca	tgccctgtaat	cccagcactt	tgggaggctg	agggtggcag	gtcacaaggt
83161	caggagtctg	agacctgacc	tttaacagtc	tattcagcct	ggcaactatg	gtgaaatccc
83221	gtctccacta	aaagtacaaa	aaattagctg	ggcgtgggtg	cacgcaccag	tagtcccagc
83281	tactcaggag	gctgaggcag	gagaatcact	tgaacttggg	aggcagaggt	tgcatgagc
83341	cgagatcaca	ccactgcaat	ccagcctggg	cgacagagcg	attctctgtc	tcagggaaga
83401	aaaaaaaaaa	aaaaacacaa	ggcccggcag	gctgagacca	tgacaggaa	gggttagagt
83461	cacctgggtg	cagtgtgggt	tgctctgttc	agcctcacta	tgaatttcac	atccagcttc
83521	tgttacttgg	agatataaaa	cagtaattgt	tggagacaaa	tcgtggatga	accaactcct
83581	gagttatggc	gtcatcaatc	tgctatgacc	agtcaaattc	gcagtataga	aacatgtctt
83641	gtagctggac	agaaaacatc	cagctctact	gctgctaacc	atcatggaat	gtacccattg
83701	gtttatccat	atagtaagtt	tttaacattat	ttttaatata	ttgttcattc	tgaatttgac
83761	agtggtctgc	caagtctccc	tgaactccac	aaattaaata	aaggtatctc	agagggcctt
83821	tccaagaac	tttaatgcc	tcttcctagc	aagctataga	aaacatttga	aaaccccaag
83881	ggcaaagtcc	cagggtgtgc	taggggcaaa	atcagtgaac	agagtagtaa	actgataaaa
83941	ccagcataga	ggcctctgtg	ggggaaaaag	acacctttct	gcagttaata	aacagtacca
84001	agaacactg	attgtcttct	cagggtttgt	agcccatctc	tgttttttat	ttttaaatct
84061	cacctgttat	ttttgttttt	agctgttcca	tagtagtctg	atggaaatag	actttatgtg
84121	ttcttgtgtg	ccaatcaaaa	ataaaataag	cacaatgaat	acctctaatt	gcatagttca
84181	tgaaaggctt	gaaaagatgc	agagcagctg	gctaattgctg	ccaatgagcc	attggctgga
84241	gccccttgta	actcagcctt	tgattctgtc	ttcagcaagc	cccagcatcc	tcagggccca
84301	tgtgatgggt	gctcagtgga	actccagact	ttgtggaggg	ctctgtgtgt	ctgctgtgtg
84361	ggaaactgtg	gtttcttgag	cctatgggaa	tgggtcagaa	agcctgggaa	tggtgggaga
84421	ataggagcag	gaacacaaa	gaggagccag	caatgcaggt	tgctattatg	tcgcatgtca
84481	cttcctacga	atatcttgat	agattagctt	ccttcaggcc	caaaaccttg	aatggacatg
84541	aaccacagta	tcagggtcaa	atgagtaaca	gcccaaacca	ggtccactct	ccagacagtt
84601	agaaaaggta	gacatctctt	gtagctggat	ggagcagggg	tctccccag	gggtggtaat
84661	tcagcaggtc	tttaagaat	gaataggatt	aaagtaagtg	aaaatggaaa	gctgaaggca
84721	gggaaggaaa	gtgaggctca	gggagaattc	tgggcaaggg	aaagttaacag	agtgattcct
84781	tagagcttgt	gagtctaaca	attttgatgt	caatggacct	tattcttgga	accacagggc
84841	aactgattga	ctttagtttt	cttttttgct	atattcttgc	cactcactca	atgactgtct
84901	attgaaaatt	gatgataatg	gccatattga	aagctctcaga	gcttatctag	gatttgagtc

84961	cctgaggcat	aatcagtg	tggaacagca	gatggatggg	gcagactatt	tttctgttcc
85021	ttcctgtgac	agttgattat	aagttataaa	aatgggcatt	ctctttgtga	tcttagcttc
85081	caaaattacg	gaaggttatc	actattttatt	attactactt	tctcagcaaa	ctcaaagaat
85141	cagagtata	ttgattat	agatttagca	atgggggact	taagctctta	taaatcaggg
85201	tcatctgaat	ctaaagatgt	atgtctttct	tatttcagtc	tgaccagtta	tttaagaatc
85261	caagatacct	ttacttttta	tttgaagttc	cttaatttga	taagagctct	accaagcagt
85321	tggttctgtt	tctctttcag	ttctctgcca	agcttttgtg	tgtcttaaga	cactgggact
85381	gggaaaagac	tgcatgtgtt	gttaaagtga	aattatactg	agccttgagg	atttgaatgc
85441	atggggagag	gtacattttg	gctttccctc	ctttgtaaaa	tgcatgtaaa	ctgatatacca
85501	gaacaatggt	gagaaaacac	caattaactg	aatgtatgaa	agacagtaaa	ataaacaagg
85561	tgggcagtc	tgtatgtgag	gcatttttaa	actctttgct	gtgtaagtgc	aggcttttgg
85621	aaagatctgc	ttttcacttt	gatttttgc	gtcccaacac	agcttgcttc	tggttctgcc
85681	ttcagggaca	tatgtgctct	ctagatctgg	gaatcttttt	gtctggagac	ctcaagattg
85741	ggatctgcac	cctccccacc	ccttctgtgc	tgaactctat	tttagatctg	cagtaaagac
85801	ttgaggcttt	ttgcaagctt	tcaactctga	ggtatttttg	aaaactgtaa	ttttgtttgt
85861	tctctgctgt	atgcatgtgg	tttcttaaaa	cacgagtgtg	attattttct	gcactccttc
85921	catagcaaac	atttccctca	aactacatgc	tccaaaagta	ggttctcact	gcacttgcca
85981	tgccgtctca	ttgcttattg	agagtgcactg	tctctttaca	acaaatcaac	tggtttacaat
86041	tttgaggctt	tttaggtggc	ttccaaactt	gtcttggggt	ttgtccctca	tatgttcagc
86101	agccaaacat	gctcaatgca	ttgtatat	cccagcaagc	catgaagcac	tattgactta
86161	ttatgaaact	taatttgatg	ataatgaggc	aggtctataa	aatagcactt	tggttagttt
86221	ttaccataa	aatagcaatt	tgtggaagac	cgtagtatg	tagcatggac	tctgaagcta
86281	gagaacttga	gtttgagtct	cagcttttct	actttttttt	tctagtaatt	ttattttatt
86341	tatgttttac	tgttaaattc	ttaacctacc	tggaatttag	tttggtatag	ggagtcatgt
86401	ttggatctag	ctgttgattc	tttttaata	taatttttag	tgattttttt	attgtgataa
86461	aatatacata	agattttatca	ttttaacctt	tttaatttag	atgtacaatt	cagtggcact
86521	aaacacaatc	acaagttttc	cacttttttag	ctgtgttaga	ggtcaaatta	ttttaccttt
86581	ctgctgtctc	gtttcctctt	ctgcaagtgg	gaataatact	cccatthaag	tgttatgaga
86641	ttaagactta	atccaggccg	ggcgtgggtg	ctcacacttg	taattccagc	actttgggag
86701	gccaaagcag	gtggatcact	tgaggtcagg	agttcgagac	cagcctggcc	aacatgggtg
86761	aacccctctc	ctactaaaaa	tacaaaaagt	taactccgta	tagtgggtga	tgccctgtag
86821	cccagctact	caggaggctg	aggcagcaga	attgtttgaa	cctgggaggt	agagtttgca
86881	gtgagctaag	attgtgcccac	tgactccag	cctgggtgac	caagcaagac	tctgtctaaa
86941	aaaaaaaaac	aaacctaatt	tcataggaaa	cacttgatga	cttgatgtgt	acaggacaaa
87001	ataagtgtct	aataatagtt	gatggccatg	aggatttgtg	gtactggcag	tccccattt
87061	gctagttaag	aagcattttt	tgagcctggt	gccatatgtc	tgccccat	atagctctgt
87121	gtggctcacag	taattgccgg	ttgaaacaca	gtggagacct	cgagaattct	cattggtaac
87181	actggataat	tttttcatca	tttttatgtc	cacttttagca	tgagactaat	atacagtttg
87241	aacttcaggg	gagaaaagaa	gatttttcta	cagatttata	tttcaaattg	tttacaagac
87301	gtatgaaat	gaaaagataa	aacctactta	ctctcaagt	tcaaggcata	atgcaagagg
87361	cttgaagagg	gtagacttct	gaatatattc	cataaacagt	atctcttttc	agaggttttg
87421	agaaaaccct	tgtttaaaaa	aaatatgagc	aacaattatt	tgttattcat	gaattctaga
87481	tgtgtaaatt	gtgtgatcaa	aatgagagga	aaaagagaag	tctcaagaga	ataagcattt
87541	tgtctacatt	taagtctttt	ggaaatgata	tgtttgaaaa	taataataccc	gttaaatata
87601	tttctttaa	ttatttagat	aaactgtttc	tattgtagtt	tctggtaaac	agtaatatga
87661	aatggtttct	attttattgt	caacaacttc	attacttttt	aatgcagcaa	aaagcttcca
87721	aatagatctg	ttaaaaagac	tgcaaaaatc	ctctgaattg	aataaaactat	ttctagcctt
87781	aacattcatg	cttcacacagt	ggaatatact	atattcagcc	accatcacaa	aaaatgtctt
87841	tcttatcagt	gctgccctag	ctaaatgtga	acttttaaaa	ttggaaaatc	aggcactttg
87901	tagacagttc	actctttggt	tcataattttt	gggctaagga	aaaaaaattg	ccaaggaag
87961	tatttctgga	aggagtctaa	cttctcaata	aacatagtag	tttagcctac	aatatttttg
88021	ttgttttggg	aattgatagt	gaccttaaat	ataacaaaag	aatattgatt	tctaaagcat
88081	ggatcgtgtt	ggttcctaaa	aagtgcagta	taatgtagca	attaaattct	gccctgtatt
88141	atacacagtt	ggtcagggtg	tcagcagaat	gccatcatca	ttattgaatt	gtgtaataaa
88201	tacttggggc	agagtgccat	agtggaaaata	aaactacatt	ttggtttagg	actctgcgc
88261	ctagcccagt	gcctatagag	ccaagccttc	acgtttcaag	aagttgtgag	gggaaagaga
88321	caggcattgg	cctcatggta	gcagctgctg	gagggttctt	ttcctagctc	attttctttc
88381	ctcatgtttc	acatttttcc	agggtgacat	cacgtgtggt	aactatttgt	atggatcttc
88441	caagcccgc	tttttagccaa	gacattttct	atgggctcca	gaccacaat	ctgggtgtca
88501	acatactctc	tacttggttg	ctcatagat	accctaagtc	aaggcatgaa	tatctaaact
88561	gatcgtgact	ctcctctctc	ccaccatcac	ctccaaacct	gtttcttttc	ctgcattttt
88621	tttttttttt	ttttgagatg	gagtcctgct	ctgtcgcctg	gctggagtgc	agtggcacga
88681	tctcagttca	ctgcaacctc	catctcccgg	gttcaagcga	ttctcctgcc	tcagcctcct
88741	gagtagctgg	gattacaggt	gtgcaccacc	atgcccggt	aatttttgtg	tttttagtag

88801	agacgggggtt	tcaccatggt	gttcaggctg	gtctcgaact	cctgacctcg	tgatctgccc
88861	gcctcagcct	cccaaagtgc	tgggattaca	ggcgtgagcc	accgcgccc	gcctttcctg
88921	catttctttt	gagcatctag	cacgtttcca	gtcattcatc	cacgtcagaa	agccccagat
88981	gccaccagct	cttctttgct	ttcaatgttt	aatggaacac	taaggctctg	gaattctacc
89041	tcacaaatgt	cttccttggc	gaatcctcct	ttctgtgcca	cccatcttca	ccttatttcg
89101	ggccttcctc	atttctcacc	tggatctttg	taaccatggc	ctctagagct	ggctccctgg
89161	atccctactg	gccatcaatc	ccatgtgcct	cagccgtaat	gccagtcaca	gttctcaaag
89221	agctattctc	ttcctccctt	cctcattcag	gctgtgcctg	gtaatctgga	gtgaattttc
89281	cactccatcc	tgcagacttc	caaatactta	tccagaggct	cagcttacag	ctcaatgcct
89341	ctttgaagcc	tttctcaacc	tttgacaggt	tgagatataa	ctatgctgct	aaatgtaaat
89401	ctaactcttc	atacattttc	tcttggttac	ctgtctcccc	ctcctgtcag	agtctgagtt
89461	acttgagggt	gtgaattgag	tcttatttat	gtctgtctct	aatatctagc	atggtgtgtc
89521	tcattgatag	cagacatcta	taaagtgtg	agtaagtgtg	ccaacgtgtg	ggtggttttt
89581	cctggggcgg	gaaggggttg	gaggggtacca	cgtagttagt	gtgagcctcc	cttcgggctc
89641	tttgactttc	cgatgagcat	gcttactggt	agtgaacttc	cttgaccctt	accagttaga
89701	tacatgctct	gggaccagag	ctgcagaata	taaaccagta	gttttgaaaa	gttattatta
89761	tctccataag	gattgggttt	tttataggca	caggtaacaa	ggttgaacac	aaccaagtga
89821	accagtgaa	ttctattttat	ttcagtgtgg	tgtctgtatg	cactgtggcc	atcccagagg
89881	ccttactttc	aattcatatc	cactaactca	agggggcttt	actgccaccc	agcactcaca
89941	cgccctgtct	cttctgcaga	gactttcccc	cagtctgttc	actgcttcct	acttttctct
90001	gcacttaaac	cttttagcata	tctccatccc	tgcctctcag	ctaaccgggc	ttgcaggcaa
90061	ttagaaagag	cttccctgtg	tacctgccac	cacatcttct	cagttgtctg	cactcgtctg
90121	ctagtctccc	ttgctgtgga	ggaattgttg	atgccctgtg	gaaggccaca	ccctccactt
90181	gaaagatccg	gtccccctcc	accttttttt	tttttttttt	ttttgagacg	gagtctcgct
90241	ttggtgcccc	ggccagagtg	cagtggcagc	atctcagctc	actgcaagct	ctgcctccca
90301	ggttctcacc	attctcctgc	ctcagctccc	cgagttagctg	ggactacaag	taccgcgacc
90361	cacgcccagc	taattgtttt	gtatttttag	tagagacggg	gtttcagcat	gttagccagg
90421	atggtctcga	tctcctgacc	tctgtatcca	cctgcctcgg	cctcccaaa	tgctgggatt
90481	acaggcgtga	gccaccgcac	ccggccccct	cccacctttt	taagaacatc	actccatcag
90541	ctttctcctt	tctcttctga	ataatcagtt	ttccccctgg	caccaagcta	tcccactcag
90601	catacaaaaca	tgctgatttc	ttcccccata	ccatgcccc	ctccagcttc	catttatctg
90661	ctctgtttta	ctttgacttg	cctgaactca	ttataactta	ttcctctcct	ttcagtctct
90721	ctcttttgaa	agttttttaa	gcttttaatt	ttggtaaagc	atatacaaca	caacatccaa
90781	cttaccatgt	caaccttttt	tttttttttt	tttttttttt	tttgagatga	agtctcgctc
90841	tgtcaccgca	cctggagtgc	aatggcagca	tctcggtcca	ctgtaacctc	tgcttcttgg
90901	gttcaagcga	ttctcctacc	tctgcctccc	gagtagctgg	aattacaggc	atgtgccaca
90961	tgcttggtca	atattttttt	gtatttttgt	agagatgggg	tttcaccatg	ttggccaggc
91021	tggtcttgaa	ctcctgacct	caagtgtacc	acctgccttg	gtctcccaaa	gtgccaggat
91081	tacaggcatg	agccactgct	cctggcccat	ctcaaccatt	cttaagtgtg	caattcggtta
91141	atgtcaggtc	caattacatt	attaagccac	caatctgcag	aatgttttca	tcttgcaaaa
91201	ctgaaattct	gtaccagcta	aacattaaca	cctcattttt	ccctcctcca	agccctgggc
91261	aaccatcctt	ataccttcta	tctctgtaaa	tttgactact	ctgggtaatc	gtccattctc
91321	tctgaacccc	accttcatca	ggcttttgct	cctacattcc	ccctccacca	cttttctcag
91381	ggctcactgt	ggcctccact	tgctggatct	gacagtcatc	tctcagtctt	tggctgattt
91441	ggcccgccag	cagctatagg	aacagtgagg	cactccctcc	tctttgaaac	actgtcttca
91501	cttgggtttc	agggaccttg	ctggcctgtg	ttttccttct	tcctcacaga	ccactccctt
91561	ttagtgttct	tctcttgggt	gcctctgtag	gttggaatgt	cccagggtctg	agtgcattgg
91621	cctcttctct	aatcttcact	cactcctctt	tagagttagt	gtctttccat	gccatttata
91681	ctcttacagt	ttcccaacatt	tgctccgagg	cccagctgtg	aacacttctt	tacctgcata
91741	tctaataagg	gtctcaacat	ttgcatctcc	agaatgactc	ttgacctgcc	ccccaatatc
91801	tgccataggc	ttccacatct	gaggatggca	actccgttct	tccaatttct	gaggccaaaa
91861	ttgggaatca	taattgacta	ctctttttct	ctcaaaccoc	catccaatca	gcaaattcag
91921	ttggctttac	tttcaaaaaga	tacaacctga	atgtgaccac	ttcttgtcac	ttccactctg
91981	ccactcaggt	ctaaccacca	tcttacattg	cttggttgtg	cacctcctcc	ctggcctccc
92041	tgccctggct	agattccaaa	cacaaatcta	gcgagaagag	cctatgaaaa	tgtgagccac
92101	cctctgctcc	aagccctcca	gtggccaggga	aatggcagac	ttttacagtg	gccgaaagct
92161	ctatgtgatt	gcctactcca	tgacctcatt	ggcctctttt	cttgattttc	tcacccttgt
92221	ccctcgtctg	gagtcccaca	gccttccctg	gtgtgcctgc	agtatgccat	cctgctgcta
92281	ccttagggct	tttctgctct	ctgtccctct	ggggaaggc	tcagccctta	aggatccctg
92341	gccttccctc	cttcttctct	tcagctctgc	actcagctgt	caccttggtt	gcgaattctt
92401	ccctgggtcac	cctgttttaa	gttgccacgc	ttctcccgca	cttcgtagtt	ttttttttct
92461	tMtattttt	ctttctagcR	cttaccaccc	tgtaacatgc	cgtatatttt	acttgtcatt
92521	tgtgtttgta	tcacccctct	ggagggtgag	ttctagagag	cgggatttcc	ttgccagcgt
92581	tcgctggatt	tgcccactgc	ttagaactat	ttctggcgca	tagtaggtac	acagcaagta

92641	ttccttgaat	gactatagag	aagaagcttt	cacaaaaacag	aaaagcgtaa	ttaattgccc
92701	acttcagcct	cctaaactgt	aggcttcatt	tgaaatttca	agttcttatt	ttatccaaaa
92761	gagaacaaag	ctgaattaaa	cattgcttca	aacgatatact	ttgagtcgtt	tccaagaagg
92821	atgtatagat	gaaattataa	tttaggttgc	tatggatata	tttatacata	taaaatattt
92881	gtgacaaata	tagtatttgt	attatataca	agaaaattac	tatggatttg	caaatcgatt
92941	catataaaat	ttactttctt	atgggacatt	aaatatatta	cagtagcttt	gaattttctg
93001	aattatattt	acttaaaaaat	gtcactggca	atacccttga	gggttacttt	ttttggataa
93061	aacttgtggc	aataattcct	cctttgactt	cttcagcaat	atgttaaaac	tggcagcgaa
93121	cctttgaata	gcttttgaat	ttttatctcg	tttgaaatta	accagttact	gatgtaatta
93181	ataccatatt	gtataaattt	tgaattctgc	atgatttcta	cttcaaaatg	atgcctgctg
93241	ccttcattct	actttatggt	taacacaagt	atttggaaaa	cctgagtagt	ggttatctgg
93301	cttagtgga	ctataattta	ctgcaggtaa	aaaggaggat	tcatttcttt	atgttgaag
93361	tctttaaaaa	tatactgata	catttccatc	ccaaattagc	aagctctgta	gaggacaat
93421	aattttatct	ttgctaactg	aagatcctac	atggcctcat	gcctaagatt	atttactaca
93481	cattgtatca	aaataaaaaa	agtttgatat	gccactaaaa	ggctttttta	aataggagag
93541	atctaataca	tcacatacgt	ggacgtgctg	tttatactac	aaaacctcta	gatcctcttt
93601	acctttgcaa	gggggggtgg	tagttttggc	tttcttatg	tttttatttg	ttttgttga
93661	agtatataat	aaaagaaaat	ctaagaaatt	gaggattatg	gttaactgaa	gtgtaataga
93721	ctgtttcacg	gtatatttga	cttcaacagc	ttccagcaaa	gaaaaatcct	ggaaaaaaac
93781	tttctgggtt	ggtcattttt	attataaaaag	gctttctatt	ttccagcaaa	gaacatttct
93841	tttatgggta	agagtttcaa	tatttttgta	attagcaact	cacagcagtg	ccttgtttca
93901	ttcttgattc	taatcctgga	atgataacat	accagcaaa	cagcagccgt	agcgtccaca
93961	gagcctactt	ctctgccttt	tttctgtctc	cctgtagaga	gaagtccact	cttttctctg
94021	atcccagtg	agtaatgtaa	atttgggcat	gcagaaagtc	aatgtgcctt	agcatccttg
94081	ggattcattt	tctttgcatt	tccgaaatta	ctcaccttgt	agcagttgtc	agatccctta
94141	attagttaact	agttttcttt	gaattatggt	atcacataga	ctgtctgtgg	tctggctggt
94201	aagagactgt	taccagctgc	tcagcttttag	tcaggcaata	aggtatttgt	tacacttcaa
94261	ataatcatgt	gatttagtga	aaggccattt	gagctctaaa	gcctcacagt	ggaaagtttc
94321	tgagacctct	gccctccctc	tcacacccca	gagaatccct	agtctcctgc	cttctctca
94381	tgctatctgc	atgattgctg	ttggcctgct	cggccaggaa	gatgggtccga	cacacacccc
94441	tgttcaactgc	ccagggtctc	aggatccctg	tcttaggcag	agtcccaagc	agtgtgtgac
94501	ctgccaggaa	gccctttggg	ctcttctgca	cccttgggga	ccagtgtccc	tacagtaatc
94561	tgtactcatg	tttattccca	tacttctctc	tagaatttaa	gtcatgagct	gtgtgtctcc
94621	ctttctgcac	tcagagagta	tcagaggagg	cctgggtgtcc	tagttatatt	tcatttttca
94681	actgtcattt	gttgaactcc	agacactgct	aactgctttg	catgcatcat	cttacttctt
94741	cttcccagtg	tttataaggt	gcaactgaac	caagccatat	agttggcact	gtgatgagtg
94801	agggacacaga	tacattgatc	attgaatgga	gctagagaaa	actgttaagt	aagtgtacta
94861	ctaatttagt	ggaaagagaa	accttcaaac	taacttacag	attgttctact	tattgtacaa
94921	aactattatt	tatttatttg	tgtctcatgg	aaccacagtaa	tagctaacat	aatcatgata
94981	gcagttattc	ttagagtgtc	tcactctgtg	gcaggcactg	ctctaagggc	tttatatgta
95041	tcattcccatt	cagtcccttc	actaattcaa	tggagccagt	actattaatg	tcctcatggt
95101	acagatgcag	aaattgaggc	atagagaaat	taagtaatta	cttccaagtt	catgcagcta
95161	agtgggtggga	ccagagtttg	aactcaagct	gtctgacatg	aggacctgga	accagcctgg
95221	gtttccagca	agataggaca	gtgcccattc	tgggctgggt	ggaaacatgg	cactgaagac
95281	atgggctctt	gtttctggag	aaaatatagg	ggacctggtc	agagctcact	tacttttagt
95341	aataaatatt	caattcaacg	aattatttatt	aagctcttac	tgtacaacaa	tgaacaaagc
95401	tcctaattctt	catggagctt	acgttctagg	aagggaagag	agaaaaattca	catctatcta
95461	catgactaaa	tcgcatlgct	atgttagaag	gttattttatg	tgctaagata	aaaagaaaaat
95521	gtagaacaga	ggaagaaggt	taagatgggt	atgaaagaag	ttttaaggag	atgatcaggg
95581	aagacctcat	ggttcatatt	ggagaattgt	tgatcacatt	tctgttgttt	gacctgactt
95641	ctttgttgat	ctcttaagag	atgcaggctg	agatcttgac	tcactcaaaa	cttgtaggga
95701	actgcttaat	atttctgcct	atttaaccct	ggaagcgttg	ggttgttttg	tctgtaggag
95761	caggggtgag	tcagcccatt	ggttctgctt	tattggctgg	gtagcactcc	cttactggRg
95821	ggcagttctg	aatgtgtgcc	ccatgggtgc	tgagttatgg	gctgggtggg	gtgtctagag
95881	gagtggggaag	tgagggtcag	acaccgagga	cattgctggg	ctgagagagg	aagttcagga
95941	tggatagaaa	gatgtagctg	cgattcttct	agagagacc	ttgcctgcct	aggtgaagat
96001	agggatgaat	tgtccaaaga	gggcacatga	ccctagatcc	tacctctacc	caaggcatat
96061	cttcaccaag	gcagtgggtcc	caggaagagt	agcaagaggg	tggagataac	agagcagccc
96121	cctcagtact	gggaaaaacag	caggttccat	gcacttttgt	ggtggggccag	ggcttttagat
96181	gagtgaaatga	cagattcacc	ttttacccag	ttattgagac	atcatgagga	gtgttaagat
96241	tctgaaagta	aatttccacg	ataacatggt	tatgRtgtca	tttggctctc	ctctgtggaa
96301	ctggtaatga	ttttgacttt	caagatgaga	tcttctcttc	tgaccataaa	acaagatggt

96361 gcaaacgagg ttcccagctg tcttgggtctg tcggcatggc tgctcgaagg ctctgggtggc
96421 tgataactca ggatgttggt gctggagagt ggggaactcg tccagatcct cctgagacac
96481 cacttttcctt ttccatacaa

ERG genomic sequence (SEQ ID NO: 4)

>21:38783451-38882000

1 tgacagattc taaagtgttg tttcaacagt ttcttcttcc tcaaagggtga gcatcatgtg
61 cacactgtca aaattaaaaa aaaaaaaaaa aagacccaat aaaacccgat gaaacccaaa
121 gagcatggag atgagaaact cactgctggc tgacgatctg aatgaatggc ctggggggga
181 atgtgaatgt ggaggtagca gaagggtcct gtcttcagtg cttgattctg Watcaatact
241 tggtaaatTT aaaaaacaaa aaaaccagca agggagattt taaatcatca aggcagaagg
301 atttaaaaag tggaaatgtg taaatgcaag agtgcgggat ctttgcataa agttattata
361 ttttcaaaaa ttcgacaagc attgttttta ccaaacgtgt ttaaagtaac agccctgagg
421 ttaatgttat gattggattc tgatgtggca ctttaaaaat aaatcatttt gcttggccct
481 agcatttctt ttcaaaatgt cctgttactt atctgtgtac aacagcagat taccgttctt
541 cccagagttt aattccatga acaccctgct gaaagacttg ttaaggcctt atggaaaact
601 ctgacctgcy tgttcattat tgacgcccac gacgagtga aggaagtga tacttgggat
661 gaggaatatt aaatcagcag ctgatctgga atctggaagc actgcccctt ggatctcgcc
721 aaataggaa agaattcttg tcaactgagca aatatagaat tcagtgttag agttttacct
781 cttgaatttt ctagaacttt gttttgatta gccacctaga gtataagaat tgctgcctat
841 gggatttttc ccagaagagt agataacata aagtaaccag cRataataat tttcattatc
901 caattatttc atgtgtctcc ataacactta atacatcatg tttctctgca cagtaattcM
961 ctatgtacag ctagggtgaa tagcaggaaa aaagagatca gatgataaag agaagggaga
1021 cagatttgtg tgacttgtat aagatacttc gcctctctaa tcctttgtcc ttatctgtaa
1081 gaaagcagca gtacccaggc tgctgtgagg atcacatgag acattatgag agtttagcac
1141 aggacttggc acataacaag cagttattag gatgatgta gtttggatta tgaMttcaag
1201 atactacagc tacatttttt atactttctt aaacatcaat ttaagtcaat ttaattaac
1261 gttattttac taggagaaat gtctctaaaa acataagaat tctatttgaa gtctacaatg
1321 ataattacat tactctgata tcttaacttc ctctttatga tttcaataa ttaaaattta
1381 agtttatgtt ttcaacaagt gcttattgaa cctctactct ttgccagtca ctgttagga
1441 cacagaagaa atggataaga tatcactaaa aacagcttct gcgatatcag tgggtatata
1501 taaattattc acactcaaca acaaacacaR gcaaagctgg gagaatgtgc tgacaggaaa
1561 gaaaatgaag agtgggagag tggacagatg ggaatcgatg tgaaatttta taaaacatgt
1621 cttggccaag ctgagtaggg gaaatgttg gaaagtaagc tagaccaaac agRgaataaa
1681 agtataaata agaaaaaaa tccagagggtc acagtggacc actagccaaa atgggttcac
1741 agcaaacaaa caaaaaaatg taatgaaata ctaaaattaa agaagtacaa ataaaaatga
1801 gattctgttc ttatctgtca gaatggcaaa gattcaaaga gtaaaatatt cagagttgtt
1861 atgggtgtag tgtggaactg gtctcatatg tggctagaaa aaggggagaac tgggtataatc
1921 cttccctcct ataactgtgt aattattttt taaattggca aatattttaa tttggcaaa
1981 acttaaaaaa tatatatacc cattggctca ataattcctt tctaggcacc aatacataa
2041 aaaagtattc aaaaaattaa atagatgggt tactgcaata taatgagtga aacatYgcaa
2101 gcagcatcac tttatacact aggatattgg ttataaaatg gtaatacata ttttagatag
2161 aatataatga aacaaaaatt ttgcttatca aaaaatgtta gtaacatggg aaaatgcttg
2221 taaaacaatg ctaactgaga agagcaggac aggtaggtag atacaaggcc aaacagatac
2281 acacRgtgga gagacatata ccaattgtta ggaatttgct gtggttggtt aggatctgaa
2341 caatttatat tgactttgtg atgactttta aacgttttct ctaccataag aattttttat
2401 gtttataaatt tatgtctaaag aacaagatgt attttttgta attcaggagt aaagaggaga
2461 gtgggggtatt ccagggtgct gttcctgccc taaccaggag cacgatgcag tcataggcct
2521 ctgccctcct ccagcagcct cggctcagaa caggacgaag ccaaccctct ctgaccgca
2581 ccacctcggg cgacccctcg acgctccacc tgcccctgtg tgtggtacct gccctgtac
2641 acctttaaca agcaccctcg caagctccct cctccttctt tcctgtcctt gtggctgccc
2701 agaccagat gtaccctgcc tgggtgtggg aggccgcaca tcagagcagt ctatggctgt
2761 gggcttcaca gctgaagctg ggccgtggga aatgttgaga cgagtgaagt cagagcctga
2821 aagacaatc ggtagaaaag gttacatccg cggggggcg tctgatggca tctgccattt
2881 attgaggatg ggtgcgagac agggtagcca ttatgatccc attttacagg tgacgaaact
2941 gagatcagac aggctagcta ccttgcccag ggtcacaaca ggaaggaaagc aaatgcaagt
3001 gtgtctcgcg ccacgtaagg tgagagaagc gctatcaggg gtggtctgac agaatcccca
3061 ggccagggag cagttgtcct ctatcctgtg ggggtcaaag aggggacaag aaggctctcc
3121 ttgaggtagc tccaaagaaa agaccccgcc agagcacagt gtgctaccaac acgctggaac

3181	agcctggaga	ctgtccctgt	ctcttctctg	aggtttctca	tttcagagaa	actctcatcc
3241	gtctggcttc	cccatggccc	ttactctgtg	attctcttat	gtatcagctg	ttaaaccaat
3301	aaatccgcc	aattgttggg	atgcctataa	tggctttggg	tctccactcc	cacagaggga
3361	atcttattct	aaagttctaa	atgtatcagt	catattttcc	tgaagacaac	agaaatgggt
3421	gtttcattta	cttattttta	tgtatgcttt	ttcaagagtt	gttaacctgt	gtttgtgttt
3481	aaataaagat	ctgttttgca	aaacaaaaac	tttatgagat	aatcacttat	aaacgtagat
3541	aattaagaaa	tggttcaggc	tcagcacatt	gtctgagagc	attgcagtcc	tgcaactgtat
3601	ttcatggggg	aagaaaagcc	cacagatgga	gccattctca	atcctgcttt	ccaggggtcac
3661	tgagtcaggg	tgccctggcat	gcctgggtgc	atgcacacac	aggcagaggg	ccagggagaa
3721	ctgtgctgtg	tccacagagg	gcagcctgga	tcacagagct	gaccaagaac	ggaagcagca
3781	agttgtttta	aagtagttgc	ctgtaaagcc	acttttcggc	aaggacaaaac	actcagagca
3841	tgctctgaaa	tgacttatgc	acagcagggc	agcggcattt	tgcccctggc	cttccttcca
3901	ttctcccggg	atccccctaa	agtaacaagg	actgacgcaa	ttcgtatttc	acttagccaa
3961	caggttgatg	aaaaatcata	aatctgtagc	tattaaactc	attaatcaca	tccatcaaaa
4021	ttctgtgagg	agttaaactt	cttcatataa	ctcttagtct	taattaagtt	tccctcacat
4081	gtgatagcaa	acgttcaaga	agtgc aaatg	tgcaagctgt	gcgtatataa	tgcaagcttt
4141	atgtctttta	attaataaaa	ttgagtccta	tcaggtcgtg	tttcaattac	catcaaacct
4201	tccaacctct	gttaattcaa	aaaaaagtaa	cttctttagg	caagtcatgg	aatagcact
4261	aggggtgccct	caggtctcca	ccaagccttg	ttagtaactt	tccatgtcct	gtctgatttt
4321	taagatgctt	tcccctgatg	ctgctggcgg	ctgctggctc	tctccacatc	cttcaattgc
4381	agatctctgt	gatgatggaa	aatccccac	attccctttt	tacacattaa	gtatctttct
4441	aggatcagaa	cttgcaacct	ggggttcagt	ttctccttaa	aagggtccggg	gagaggactg
4501	tatttctgtt	gtaattctat	gtatttcttt	ttaggcacca	gaaagtatta	ttctgagaaa
4561	ggactgtagg	cttcaactaga	cttcccaaat	tgcaagtgcc	attaaaaaaa	aaggaatcag
4621	gaaagtcttt	gggaaattag	gctaaaactc	ctaactgaag	cacaatcgag	ttttcagtgc
4681	aacagcagga	tccctgcagt	ctgtagcaaa	cactgagatg	tttgtgggga	cagtggggcc
4741	ggtagaggtt	agcacctctc	aatattgtca	gatgtcacca	atgcccagg	gaagcaccaa
4801	aaccaatcag	tgaattgtct	gcaaacatc	agctttgttg	cttccaagtc	tttactagaa
4861	ctcttaagct	gatcttaaaa	ttcacagaca	cttagagata	aatgaacact	ctcacataag
4921	ttcccaaaat	tcttttactt	cttccaattg	aaacattttg	ttcccaattt	caagttacca
4981	agggaaaatg	atggtaagaa	tcacaaaagc	tggtatcgaa	tatttagctt	tctatagtgt
5041	ctcatagggt	cttacaatg	ttaatgcaaa	aaaataaaaa	ctccactcaa	atgcagaaaa
5101	ccaacttaaa	aaattcagag	catctgatac	ctaactgact	aattgattta	tttactgaaa
5161	caacaaaaaa	acagagtgc	ttgagactgg	aaattgcctc	tatggggggc	tgtttcctaa
5221	ttccagtggg	cagccatcat	ttccctatg	ctaacttggg	ctttgcgcta	acgtaattga
5281	aactgtagtc	aacagactca	ggactcctt	tttccactct	ataaacaagg	ctgtaaaatc
5341	atctcattac	tttaccagg	aaatctaaca	gaaacatgtc	ttaaaggaag	tcacctataa
5401	agagggccta	ttttaacaga	agattaaaaa	aaaattcacc	tgtagtcatt	ctaagggttc
5461	acttgggtac	atacaataaa	cacataatgc	aatgtaataa	aatgttaaat	caagcatagt
5521	accttagaca	agtgtaaact	attatttcca	aaaaagtgtt	cagaaaaatc	caaaatattc
5581	aaactttcaa	aatagcttca	tgtattcata	aatgtatact	tcaaatataa	cagtaataca
5641	catttgggatg	ttttgcaaaa	tgtacctacc	ttctaaagga	tggtctgtta	tttcaatgtc
5701	ctttttaatg	cactgagttt	tttgggaaaa	aaaaccacat	aagatcgtag	ccaagctatt
5761	caagattttac	acagaatatt	tgatttaaca	tgcatatgaa	taaggatgaa	aataagaaca
5821	aaaaataaac	attttgggtc	acaaaaatca	gtaatttcta	ttcatgtagc	taYaagtagg
5881	gcacattaac	tctcaatttt	ataattcctg	tctataggag	ccatatttca	gcagattcca
5941	atccttgaag	ccaagaaaag	aaaagtataa	ggtgaaagtt	cttggtacac	agattttttaa
6001	tatattattg	aacacacctt	gagaaataaa	agctgaactt	ctcaaacaca	gtttctatgg
6061	tacatatgag	tcaaatcttc	taatacaaac	atttcaaatt	atgaaaaatac	caaaccaaaag
6121	agtatttttg	aaaatggctt	ggcccaattc	tgggcaaact	tccattttct	tatgcacttt
6181	ccaacaaact	aaaagcctta	tgaaaaacac	accaaagctt	tctcacctcc	gtgcagcaaa
6241	actaaagaca	tgcaacaaac	acaattttct	tttccctcca	ctcagcatct	gcttttgttg
6301	ctgatttttc	acatttctac	aaatgtcacg	aaggcatggt	ggcctttaaa	acaactgaac
6361	ggacccccat	tcaagactgc	atgccccctg	acttgtagtc	accattaaat	tggcttcatt
6421	tccaggaatc	aggccatatt	taggattgta	cctgtgcaga	tttacctcca	cattaatctc
6481	tacatgctat	ctactaaaaa	cttaggcaag	gaaatgcatac	agacaaaaca	ccccacagca
6541	cagagaaccg	accggccatt	gctttccaat	ctccgcaaac	ctaaccattg	ctggaagaaa
6601	tcttactcac	agtgcacaga	cagtaggtat	tttattgaag	ataaacatat	agtggaaaca
6661	accaaattac	ccccatttga	gttacgtgag	cactcagttc	tcagcgtgga	tgtcccacaa
6721	atcgaagtcaa	catttgcgtc	ccattaccag	cagccacttg	ccgagtatct	cttcgcttcc
6781	actgggactg	cctggcatcc	ctgatgctaa	ggagccactg	aagagcctcc	aatgtctgta
6841	cattcacaaa	cgcactcttt	gctttgacct	gaccttcaa	cctctccgag	tctgtgctct
6901	tttctcagac	acacatccag	gcaccgttag	ggatagttag	agaatctgaa	aattcagaag
6961	cgctccgaaa	agcctttcca	aaagtaatcc	acagcactca	acagtgaatt	tagaaacccc

7021	aatttttttc	tgagtttgaa	gtttttaagc	cttgcgatg	gttgagtag	gaaaaaggaa
7081	atttactagg	cagtgcagg	gaaatcttgt	tgtcctctat	tgtggcagtg	ggggtgttgc
7141	ccaaccctaa	cttatctgcc	ttgataaagg	aaaccacaga	aaagagtaac	aagaacaaga
7201	ttttgtcaaa	ttaaaaggaa	ccctttcctt	accttaatat	tgctggccat	aatgcRatca
7261	agttttattga	tcgttaataa	atgttaataa	taattattgc	ttctctctga	ccagaaagta
7321	gttttgatga	ggttgtttag	agcggtatgag	attgtgctaa	gtctgggaaa	tgaagtcagc
7381	caatggcagg	aagaggtttc	tattggctct	ggctgtccag	cccaaagaaa	caggatattt
7441	gggagtgagg	agataagaga	ccctgaaaac	aatgttggtt	ttcttgatga	tatgcagcca
7501	ggagattttt	tttttttaat	taaaaaaaga	aaaggcatca	attgggatgg	ggactgccac
7561	agcaggtgtg	accggtgtgc	cgccgtgtga	cacactgcac	tgagaccaag	gcaggatgca
7621	gatgtgatgg	gactccgcat	ggcttcacac	gggctgcaag	caccttgagg	ccaaggcggt
7681	gagggcacc	cactgccctg	ggtgtcagcc	cttcgcagcc	caattcttcg	cagaattact
7741	aggacagagg	acttgagctc	ctttctccta	aaaggaaact	ttgcaggtgg	agtttatttc
7801	atgttaatat	atggccatgt	tcagtaacag	ccattgcctg	gctgattttt	aacaacctat
7861	atttattcaa	catttcatat	aagtgttcca	gaacagtttc	attttctcct	tccaatatcc
7921	tgactttttt	atttgctcta	caacaaagtt	gttgaaaacg	caaggactct	aggcttacag
7981	taaacacRaa	aaataaagag	gaaaaataaa	ccttcctaag	tcttgttttc	aagtatttat
8041	taaaacccaa	ataactgaag	tgactacaaa	tgtcccgga	tatcagttag	ctggtctcac
8101	tctgacagac	atccatggtg	cagacaacag	atcctcatag	aacttttggc	taccacagag
8161	tgccatgtga	gggccccatt	acatgtctaa	aatccaagct	ataagttcag	ggtcacaggt
8221	cKgtttttctc	ctcgaaggaa	gtacagcgaa	tgaggccctg	aacattcctt	agagggtttc
8281	aggactttta	ttactcattt	tcaggaaatt	gcttgaaga	aacattttgc	tttgaatcta
8341	gctacaggaa	cgcaggaccc	ataaagaggt	gtggtctcaa	Yaggcgaccc	caagaacaca
8401	atagactaaa	tcctgagtca	tctgacaatt	cctggttgca	gagctggacg	ttcagtaaat
8461	ggattttcact	cagactttag	gcggcctagt	gtcagagttc	tgtccaccag	cccagctcat
8521	tctggctttt	attatatacc	tctgatttat	cacctcatgt	taggacaaaa	gaaggaggag
8581	aaggagaaag	ataagggggg	aaagggagag	gaaggagagg	aggagaggga	gggagggaag
8641	agggcgagaa	aaaatggaga	aggagaaagg	aaggacagaa	gagagagaga	aaggaaggag
8701	gatgggagga	aggagggagg	aaggaaggag	aaagggaggg	agggagggaa	agaaagaaag
8761	aaaagaaaa	aaagagagta	tatttgtaaga	aaatcattct	gtggaaatca	gaaYctaagt
8821	tctcagcacc	tacattatga	gggggtgttg	atttccgctc	ttttcatctc	tgaacaatga
8881	ataagatggc	aggctgatta	aaattctggt	tcccNgaaaa	tttcaaagtc	ctgagctgtt
8941	ttatctgggc	agcttccact	agaattctgg	agtgcggagg	agaaaagctc	tctcagcttc
9001	cctgagtgtc	cttgcttttt	gttctctcct	aagaagcatc	aatgtaaaat	gttaactgtg
9061	gcctccacaa	cacatggcac	ctgattatgc	ctttaccaaa	caccagcttt	aaataagatg
9121	aacgctttgc	taatgaaata	gccacggaag	aaaatcctgt	gtggtcccgc	ctcaccacag
9181	cctcgtggtt	gctaattccag	ccctcagttg	ttccctgcag	caggaaagat	cagcatttct
9241	actgggacct	aatgcttcag	tgatgatttg	gccttaactc	cctggttctt	gcctaatgca
9301	aatcaaatga	tgaagagtgg	gatttctttac	ctctctattt	tgaataaatt	tccaacttac
9361	agaaaagcca	caagataata	caccaaactc	ctaataccct	ttgcctgggt	aactctgatt
9421	gtctactggt	gccaatgtac	atcattttct	ctgctttctc	tcacacatgc	tctctctctc
9481	tcaatctctg	tctctctctc	ccatatttat	ttatgtataa	catacatgca	tgcatatatc
9541	catcattttc	taactaataa	aaatatttgt	ttataagaga	caaatttatt	tattggatag
9601	taaaaagag	aaaaacagat	gctattattt	gatcaatatt	cRatccataa	gttttctgct
9661	gtctaatttt	ggaaatgcat	acaaacataa	aaagtgtatc	aataaaaaag	taatttgagg
9721	atactgagtg	gtaaattctg	ttccaaaatt	ttacaaagga	aactcctcct	cattttctaat
9781	gacaaaatgt	ttgattgatt	cctggtgtct	tcaaagagaa	taattcttga	tgtttcaaatt
9841	ggcatcatag	catccaagcc	tggtaatctc	tgactattta	atgtaaaagt	ttttaaaaaa
9901	aWcttttggt	cttgatagga	caaggcacat	cttttttttt	tttttttttt	tttttgggac
9961	ggaatcttgc	tgctgcccag	gctggagtgc	agtggcgaga	tctcggtcca	ctgcaagctc
10021	cgctcccag	cttcacgcca	ttctcctgcc	tcagcctccc	gagtagctgg	gactgcaggc
10081	gcccggccacc	acacctggct	aatttttttt	gtatttttag	tagagacggg	gtttcaccat
10141	attagccagg	atggtctcga	tctctctgac	tcattgatcca	cccgcctcag	cctcccaaag
10201	tgtgtggatt	aaagacgtga	accaccagc	cggccaagg	cacatctatt	aactagaata
10261	tagcagtgcga	gaaacggtt	cccaaattat	atcctatggg	gaaattatct	gaagatatag
10321	atatcactct	agctttctct	acttgtaccc	caagtctatg	ttattaaatg	gcctcatccc
10381	tacaccaccc	agcagcctaa	gaaccaggct	gcctttccca	cctgatacct	cctctcccta
10441	gggaaatatg	ccctccaagc	caccacggcg	accaagttag	ggggcaatct	ggcacacggc
10501	ccctacctga	ggccgctcct	gccatgagcc	tatcagggca	gggtcctttt	tactctcct
10561	tgagatgcga	taatcaccac	actagaccag	cacagcccat	acagtgcctg	tggacagcRg
10621	aacctatgaa	taaagcacac	atttcccttg	catgcttctc	tgccacaaac	ccatcccctg
10681	tgaatctgcc	aaaaccttcc	actagtctcc	tcacatttta	ctccaaagaa	agaaacaaaa
10741	taatataaca	taaatatacc	ataaattcac	tgtcacataa	aaagaaacac	tagtatttat
10801	ttatcctaaa	atcattttct	cccattttct	aaagatcgaa	caagttctca	ttcttgttcc

10861	cttactaact	tctctgtatg	gaatatgggg	aggaagggtta	ggagctaaga	acactctgaa
10921	tgacagYcac	atggttctca	atgtaataca	tgcgagagc	aaagagagga	gatccgtctg
10981	aacctccaca	tctgtgggga	cacaacagaa	gattcactca	tgactcattc	aYgtctttat
11041	ccctgcccta	tacatcccta	caatggaaacg	cccttcctcc	tcaaaaaacat	tgaggaaaaa
11101	tgatgtcgta	gatgaacact	tagcaatatg	gaaaaactgct	ctgatatttt	acctgggttag
11161	ttaaaaaatc	acagtatata	ggagaattat	atccatagac	tcattttgtt	aaaaataaaa
11221	aggcagttgg	ggatggccgt	gcattctgtt	ccaaagcatc	tgtgaacctg	aggaatagac
11281	accatgagcg	aagctaacct	tcccaggttg	aaaaatgaaa	tgcaatggtg	gcagacatgt
11341	ccaaggaata	ttatggggac	tcgatccctt	tatgaatctt	gtcatagatg	aatgtgtgga
11401	gatgtcaact	agttggcagc	agaacagtat	tggaacggtg	gtaatacaag	gaaatagtat
11461	catcatatta	gaaactttgg	aatgagtata	aataatggct	gttcaacaga	aaaacccatg
11521	tcccctctcc	aaagggcctg	tttcaactata	tgtaaaaaat	aggctcatgta	tgttttcata
11581	ttagactttt	tgtaaataaa	cctttttttt	tttttttttt	gagtctcact	ctgtcgccca
11641	ggctggagtg	cagtggcgtg	atctcagctc	actgcaagct	ccgcctcctg	ggctcatgcc
11701	attctcctgc	ctcagcctcc	tgagttagctg	ggactacagg	tgcccgcac	caccttggtc
11761	atttttttgt	attttttagta	gagacaggtt	ttcaccatgt	tggccaggat	ggtctcgatc
11821	tcttgacctt	gtgatccact	cgctcagcc	tcccaaagtg	ctgggattac	aggcgtgagc
11881	caccacgccc	ggccaataac	cttttgaat	agtcaaaaaa	taaaaataaa	ataaaaaggc
11941	aaagtaaaaa	tattgacagc	tatatcttat	aataccaatg	gcagagaagc	tttttgtttt
12001	ttcccctctt	cttctttggt	tcacttatat	tttcaaaaac	taaatttgta	ataaacataa
12061	attatatcct	aataataaaa	tatgtaagta	acatgcaatg	ttgttcattt	ttgtaaatat
12121	ctgaaaaata	caggcatggt	ctgtaaaaag	tgagccccc	aatacttgta	ggtgtcaaac
12181	cttttagcat	gcccacatgc	ctgttctttt	ttaaatgtct	acagcaagca	tagagctctt
12241	tttttttttc	atttttattaa	gaaaaaccag	aaacgctcaa	gagtttcgtg	gaggccatct
12301	atgoggcatt	aatccattca	ggtgatatta	atggcctttt	cctgccagga	actccagtgg
12361	gcactcagga	agccaggatt	acagggtctc	atccagtcac	cgcatgtctc	ggcattcggg
12421	aagcctatct	cagtctcctc	aaattctgca	cacataaaac	ttcagagcct	gggagcgacc
12481	catgaaacgc	agggtttttaa	ggacaaaaca	ataggacaaa	agtgtctgtat	attgtcctta
12541	agatgacaca	tgaattttaa	atgcatagtg	tttggtattat	tcatgagagc	cccgaagaa
12601	cgatgtccca	ggggtttttg	ccaggacatc	gcctgagatt	gtagataaga	atcaagcatc
12661	tggggctgcc	gcacaatgga	aaactccagc	actccatgga	acttttccat	ctcgacagt
12721	cggaggattt	gcctgagaac	atacgcggca	tgaagacact	ctcagccctc	aagggcaccc
12781	agtcagcgct	gtttaaggac	ggtttttctg	ttcacagcat	acttcattgat	tacagcttca
12841	ctccattgga	tacaactatg	tgagtatgtg	tgtatgagtg	tgtgagtggt	tgtatgaggg
12901	tgtgtgtgtg	ggggggtagg	ggagtgcatt	atccccagc	agcacattaa	gaaataatcc
12961	aattacaatt	taataaacca	ttgtttcaac	actttctcta	agtgggtgaat	gtattcttca
13021	gtctcttggt	tgatctgaac	taatagaaac	caaggaaact	gttatcacat	accaaatacct
13081	caacttctca	acaccaagtt	gcaatttctt	taatactgaa	gcacacgtca	gttcagttct
13141	ataagctccc	atggagcaat	gatttcaata	aacacaattt	cataagcatc	ccacacgttc
13201	tcaagtttag	gccaattttg	ctcttactca	actccatag	actttaaaaa	tgagggaata
13261	ttaaaacat	taccaaggac	tccaaatagg	tcaatgtatt	atctgtggtt	ttagaataaga
13321	aacctaagag	gagagtaatt	attttgtcac	cattatttag	cttaaaaaaa	cttttaaatg
13381	tattgttcca	attattaggc	ttctccatat	aatttggaag	ccattaaatg	agtttcaact
13441	tctttgtccg	tataactggg	tatatctgag	catagatgct	acagacatca	cgttgccatt
13501	ggtgcccata	aggcttcgta	tgcccacagg	gcataataat	aagattcaca	aaagcatctg
13561	actggcatcc	cactctaccc	ccgactcaga	tccaaagtat	cctttcccag	gtactgtccc
13621	ctgtcaccat	ccttggccaa	attgagaatt	tatcccagg	tgttgaaatt	aattttgatt
13681	ttgattctag	tcagagtttt	aaatgatttt	ttagaagtgt	tgtcaatctc	acctataaga
13741	tactaaacag	catcacttat	ttttgtctta	atctgtcact	attttatgaa	ttatttttaa
13801	agaaaatata	gattacttct	taaagaaaga	tcataagtgg	cactataata	gcattcaatt
13861	gatagaaatt	gatgtagaag	cctgcattaa	taatttttcc	tgctgtttc	taaagtggc
13921	tttttagttt	cggttgaggt	gggttttaat	tcagcatcac	ctcatttgat	gattttattta
13981	ctaattttatt	tccatcacaa	atggaatcag	aaatgaagcg	aaggcaaaaa	atctttgcc
14041	caaatttcatt	tatatattgca	tgagtatcac	caggctacct	gcggatgcaa	tgaagcccta
14101	taccctgtgc	tgaagtcagg	agcacaggac	actgggagtc	tgtctcagcg	cagtaatctg
14161	tacatggtac	gtccacatg	gtatctcttc	cacctgagcc	ctcctcatcc	ctgccccct
14221	ctccacacct	cacctctctc	acactgtatt	atctcccact	gcaagcaacg	cagaggacac
14281	agaggtgcag	ctgaagtga	ggcaggccag	ttcttttcca	attcactctt	atgaaaatgt
14341	gttaaattccc	gttaaattgca	ccgacctgaa	ctgaaaYgat	gatatacagta	aatttgttgt
14401	tgtcatctga	tgcccacaca	gatctcagca	catgttttgg	gaaaggctgc	agtgttcaag
14461	gagtactaga	aatgtttctt	ctagtggac	tggccatgat	aagccaaaga	atatattagc
14521	cgatagggat	aaatcatggt	tttcccactg	cgcggttgcc	tggtggagat	tattccagcc
14581	tttattttaca	gctacaaaaa	taatcctgcc	tcaaggattt	taggaagcaa	ccacctcctg
14641	cctaagaac	tacactgcat	tgttgacttg	ccagctagaa	accagacatg	agtcgtgacg

14701	ggaagttccc	tttgaacggc	ttgaattggt	gtcactagag	ggtgctaaat	gcccccaaa
14761	aaggccttta	gagagaatac	tagtgtgcct	aattcttgat	ttaaattcct	tgaattgaca
14821	ttattttaca	aagtgggtctc	attttttact	tgttaaaatt	aagatctaac	atttataatg
14881	cagaaaatgt	gttttttttt	gaaaaatttg	gattgttcat	tattatagaa	cataaagaaa
14941	tacacgtttt	aggagcaaac	atgaattaca	cagatctgaa	tataagatta	aaggcataaa
15001	tctggggggg	tgagatttgt	ccatggaaat	tgtataattg	ttttatgcca	ctgttaatga
15061	ttttaaaagg	tgaaaaattt	accactttac	aataatttta	gattgttcaa	tatgttccca
15121	aatgtgagtt	gcatcaaatt	cggtaaatgt	agagtatggt	tgttttctct	gtacacaata
15181	ctgacgtcag	tgctagtgcc	tgccataaact	cagacaccag	tagatccttt	tccagaacta
15241	agtgttatag	gaggaatatg	gattatactg	taataataca	gatgtgaaaa	aagcctcacg
15301	gtcccttcct	ttctcaaaaa	aaacgtgtag	tattttggaa	tatgcatgga	gtatagcact
15361	tctaaaaatg	gtacttttatg	tatatatgaa	gtaaaaatagg	ccttaaaact	taatatacag
15421	atcttcctga	aaataatcct	attatacatt	tacactgact	ttttgggtcat	ccaaacatta
15481	tgactttaag	caaaaagaaa	tataaccaat	tttgcaactt	tcctcttcct	gcctttgctt
15541	ctacttaaaa	atatctgcag	tatttccttt	tctcgcgcgt	tccactttgg	ttcatatacc
15601	aaggaatYgt	gaggtcaatg	ggtgcacagt	gtttgtctgt	gttcacggct	gtcggggcggg
15661	gctgggtgtg	ggtaaggctc	tgaccgacat	ctgaagaggc	ttcacccact	gcggttttagg
15721	atacattttc	ctctcgctct	gccagtagaa	cagggaggct	agttttcaat	gtgttctaca
15781	gaaggataat	taaggcagtg	tgtcacctgc	tcgtctgcta	tttttaaaag	gaaacaaaca
15841	taacctttgc	atgtgagagg	cattgcaaac	tagatgggag	gaggcccggg	ccctgtaaac
15901	gtcgtacatt	ccgggtcccg	cttcgctSgg	gtgcaactgc	agataaaaaga	cgctggctcg
15961	acaaaaagcg	cccctgggct	ctgcattgcc	agcatgcaca	atacgaattg	ctaattatga
16021	gctggccttag	gtaggaactg	aggaccctga	gtgcagggct	aacctctgct	tcacaggagg
16081	ggtccactgg	cgcgcgctgg	aggcggggag	ccttcgtgat	atttctccga	agggaagagt
16141	ccatgcagca	tcctgataac	agggaccttg	gcttcatccc	caactgctcg	gtgaccccgga
16201	tgcgctcatt	aattcatttg	tgccaatatt	cctttttcaa	tcaaccagta	actttgcagg
16261	ttgtttcaaa	gaagaaatta	aatgtacttg	gagatRgaag	gcagccatct	tcccggagcc
16321	ccagtgtctga	caactgggagg	ggaaccccg	gattccctca	tgactgtgtc	agtgtggcc
16381	ctcaacaagc	acttactgaa	tgagtgaatg	aatgaatgaa	tgaatgaatg	aagccggcca
16441	cagagcccg	gagtatcagg	ggagagcaat	actggctgtg	tcagtggcta	gtccaggaga
16501	gcccttcaag	agcccagagt	tcgacttgaa	cttgggactc	ccctgggatg	ggtgaggttc
16561	tctcaacagg	agttgacagg	agggaactgg	gacaggccca	gccagtttca	ttcagtcag
16621	ccagctgtgg	ctcctcttta	ggaatgtggc	caggcccagc	tgtggccaaa	gcaccagcca
16681	gtccctcagt	tctgaacaga	acgtgcccg	catcagtgcc	ttcatttggt	tttaactgc
16741	ctgctgacca	tccacaaagt	taagatgaca	gctgggactt	aagttgtgag	gaaccagtgg
16801	ctcccccact	actctctctc	acacacacgc	acatgcacac	acacagagac	acacagacac
16861	acacacatgc	acacacacag	agacacacaa	acacacatgc	acacaaagac	agacacatat
16921	aaacacacat	atacacacat	gcacacacat	atatacacac	agacacacag	gcaaacacag
16981	acacacaatg	gatgggtatcg	atatctatga	acagatgaga	gagagagaca	gagcaggagg
17041	tcactctgtc	cccacccctc	ttctccccac	tgtcaccacc	tctctcccaa	gtctagctgc
17101	aggcctgtgtg	agctgtggaa	gtggcttttt	taagctgtct	tcttagattg	agtcattggg
17161	cagtgaggaa	ggtaaatttt	ctctaagaag	cctcccaaat	ctgtatctta	tgggtattcc
17221	gtctcaataa	ttgagctaata	agtacctgta	agcaagtggg	tagtcaataa	aaaaaaaaata
17281	ccggctgttaa	gaacgaatga	atgaatgagt	ggatgagtgg	atgagtatac	atcctctatt
17341	ttgctggcat	catgattatt	ttacagcaga	tttatctgga	tttcaatata	aaataagagc
17401	tcccatttgt	atacaactta	gtccttgttaa	ttttccaaag	taactctgat	taaaattatt
17461	tccagccaaa	ctattttaggc	aactgggttg	gtttcacacg	tacttatcag	aaaccattag
17521	aatgggtatat	tgattgcatt	tctagacatt	tgccatacca	ccatggcaac	tgtatttccc
17581	aagatggccc	agtaggaWgt	cccacatgc	ataaccttct	ataaagtgc	attgactttg
17641	cttccatcag	aaagtggagt	ccacgtttccc	actccttgaa	tctgagttgg	gctgtgacta
17701	cagcagaagt	tcagctatgt	ggcttccaag	gctaccttgg	aaaggtgata	gagcttccac
17761	ctgattttct	tggaacacgc	cactcttaaa	cctcagccac	cataggggga	gaaagcccaa
17821	gccatgtgac	gaggccattt	gtaggtgttc	cagtcacag	tcccagccca	cagccagcat
17881	taacccccag	aggtataaat	gaagaaacca	agatcactcc	agccccaacc	ggcacttgac
17941	tgcaaacctta	taagagactc	caagcaaata	atatctatct	cagcccaatc	aagccctgga
18001	accatgagag	attctaaata	ataaagtgat	cattgtgatt	ttaagccact	gagtcctgca
18061	gtgagttggt	atatagcaac	cgggcatcgg	cataccatca	gaagattcct	taagaagaaa
18121	cttttgcaac	gttgctacag	cagaaaatat	tccttctgta	tgatcattca	tccattcatc
18181	aatcttgatt	ggcattgcct	atgtgctaag	cattgtcggt	cactggagggt	gcagaaaaga
18241	agtgaagcag	tattcagcca	aaacaaacag	ttgcatattg	gatagagaga	aaaatatgcc
18301	aaaatttaca	cagtattttg	accacaacag	aaggacatga	gaagggaaca	tcaacagttc
18361	cttcctattc	ttttccaccc	tttttctttc	taggctcatt	aaatgtagag	gtgcctcagg
18421	gtcactttgc	cccatgagag	tagataccct	cctctcccat	ccctaacttt	ggtcttcact
18481	ctatgcctga	taccagcagc	tccccaggct	cctccatgct	gttggcacc	cattcccatt

18541	ctctggaaac	atctcagccc	atcaacctca	gattactttc	ccaaggagat	gtgagaaaa
18601	ttcaattttc	ccctttgtga	cttgacagaga	aatagggggtg	gtaaatgaga	agcagagggga
18661	gaggggtggg	attcattacc	ttgagttact	aaagagaaaa	gctctagtgt	ttagaagggtg
18721	ctgttattat	tttctctgaa	acattagctc	ctagaaatca	gaggctctgt	ccctagcatg
18781	gtctctaagg	cttccaggac	gccccagggg	cttccctcag	tcaattttga	ctctaattgag
18841	agccatctgg	aaggctgcaa	tggtgggcct	ggctgatcac	tgagggggcc	ttcgctgtgg
18901	gcacagggac	gtgcacaggt	tcacagagtg	tgtctccgtt	tctgggaacc	atggctctgt
18961	gtgctccttc	cccaaagaaa	gcaggagaaa	gaaaaaccac	agtcctgagt	gggattctcc
19021	ccaaacatct	aaaaagcttc	tttaaagaag	tcaagttatc	ataattaaac	tcacatatgg
19081	aatgcctata	gacaacgtgt	tgctcttcca	atcccaggga	gctaggacac	ccaaaactca
19141	cctcccagac	ggagctagga	tttgagatag	tggtattccac	attacttaaa	gctgttcgag
19201	cttttcatca	ggaatggaaa	gcattaatgc	gtctcacacg	tcactctctg	tttcatggaa
19261	ttcttacact	gRggagctct	ttttatttta	ctatttttct	ttaacagcca	gaacactgag
19321	aagtttgcaa	agaaatTTTT	cagctgcctt	tagtgacctt	taactcaaga	ggttttgctt
19381	aatctgtaaa	ttggagttga	aaatatTTTca	gtactttttg	agtgggggagc	tttttgtctt
19441	taaaagagtg	gaacgacact	aacaataggg	ctatacatac	cttggtgcttc	tttgcagttt
19501	actgcaaacc	aaaaccagat	gtgaagtatc	agcggctacc	tcgtagtgc	ccatggaggt
19561	tgaggctaatt	ccatatcatt	cttaataact	caaaactgcc	accttttaac	cacctaatTT
19621	ttttctcact	tacaaatgca	aaaaaaggaa	ttgacagtac	acattgaaat	cgattttccc
19681	ctgaccactt	acagattttc	tcctataaga	gccaggaata	aaagctcact	ccattataatc
19741	aagcaccat	catctgagta	tgcttttagtt	taggtgatca	tcacctttaa	caaaccaaga
19801	tctggaagga	agattttctga	tctttaacaa	accaagattt	gaaaaactaga	tattcctgtc
19861	atgacctttt	cccaacattc	agtatgtaag	ggattcactg	gattattaga	ctgcttgtaa
19921	atgtaagaaa	acatagaagg	tttagtggtga	aggagtttagc	aacctaaaca	atgtttcccc
19981	ctgaattttc	aataaccttc	ctgaatttaa	aggtaatgaa	tataaatgaa	gaacaaaatc
20041	aatagatata	ggaatgattc	acaacctccc	aaatgaaata	atcRatcatt	ccaccttaca
20101	tttgaacatt	caatcaatat	tttgtctatt	tactaaatct	aagaggaaact	actttaacat
20161	aaaactaaat	atgaaattat	ttccagcatt	gaaatattgtg	ttttgcctta	gagcaggcac
20221	aaagtgggaat	gcctatgggt	acaaagctag	gcatttgattg	atgtttaata	cctacaaata
20281	cagcaattac	actgtcatgc	tacacgagac	ttcctatttc	tacactgaaa	ttagactgca
20341	catatgtagg	gataaaagat	aagacacaa	catatcatat	ttcctcaaat	ctaagacttt
20401	gggtggtaag	atgtaaagca	catcatgatt	taaagtatca	ctaggaaaga	aaaaacaatt
20461	ctgccaatga	aaattccaccg	taatactttc	taatcaatca	attttaagat	acatcccaat
20521	ttttaagcaa	gaataaatgt	gttaaaatta	acacaaattg	gtacacttgt	gatacaWcag
20581	aagcactaaa	atgtcacagc	aatagaccac	agggtttctc	aaccttgga	gcatcaacat
20641	tttgggttgg	tttaattatct	gttgtggggg	ctgcctgtg	catggcagga	tggttgctag
20701	cattcctggc	ctctacccac	tagatgccag	taacacctcg	cagactccaa	ttatgacaac
20761	caaaactgtc	tttagacatg	gccaactgtc	cctaaagggtg	caaccaaaaa	ctgtcccagt
20821	tgagaaccac	tggaacagag	caattcggtt	ttatcagcct	tgtaataata	ttgtgtgtgt
20881	gcactgtgtg	catgacagac	tgtcagccta	atggagggtg	tattgtcaga	tgggatagat
20941	tgacctgaaa	aaatcacata	aaccaacagc	tggtacatat	acaagagtgt	gtcctgccca
21001	tcctaaccce	cacaatgcaa	aagaacctgg	atgaccacga	gtcaccatca	caggagtTTa
21061	ctatccataa	gtgtaaacca	tccaagaaaa	gtgaaaacag	attaaagaaa	ataaaaaata
21121	ataaaccacc	acacaatatt	tgggccttagc	cagataacaa	caatgaaaaa	atgtaattatt
21181	ttgaagtatt	ggcaagaatg	tgagggaaaa	agcattgaca	ttcgtgtata	atgggaataa
21241	aataatttat	cctttgtctc	cgcaattgat	attctagaaa	tctatttttt	agaaatactc
21301	acaatagaca	gttataagta	caaaggatgt	ctgctattac	actgctcagt	aacagtaaaa
21361	tatgaaaata	cactgtacgt	ctttcaatca	gatattgaat	aattatacta	cacagtatta
21421	agaatgagat	agatatatta	gtactgcata	gaaaaatattc	aaagtataat	aagttttaga
21481	aaattatcac	atatgtatat	tagcatatgc	agagaaaaca	tcttaaaata	caaatttcag
21541	ttttgaggtt	tggtcatgg	agtaactgcc	cccttttcaa	gtcaaacagt	tgagaaaatt
21601	taggataaaa	caagtgtgtg	ctacctttgt	aatcagataa	agtattttta	aataattcac
21661	attataatat	agaaatactg	aaatgccaa	aaaatttaaa	tatatataat	aaaacctacg
21721	ggccaggcgc	agtggtctac	gcctgtaatc	ccagcacttt	gggaggccga	ggcaggccga
21781	tcatgaggtc	aggagattga	gaccacctg	ggcaacatgg	tgaaaccctg	tcttactaa
21841	aaatacaaaa	attagctggg	cattggtggc	tggtcctgta	atcccagcta	ctcaggaggc
21901	tgaggcagga	gaatcgcttg	aaccaggggag	tcggagggtg	ccgtgagcca	agatcatgcc
21961	actgcactcg	agcctgacaa	cagagtaaga	ctccgtctca	aaaaaaaaaa	taataataat
22021	aggctaccaa	ttagacttta	tagaatacat	aaatatatta	atatagctac	aaagagggtg
22081	aagcttattc	attgagacaa	gagtatTTTat	tgagcatcga	ctttattaaa	tacactgtta
22141	gtaagataac	aagtgaact	tgTTaaaaaa	aaagaaacta	ggacaaaaga	aatttgagtt
22201	gagttcccca	aaaacaggga	atattcagat	tgtaggggaa	attgggtaaa	ggatttgttt
22261	cagcatgaaa	ggagcaagga	gcaaagaaac	taaattgggag	gttatgctga	gaagtatatc
22321	agtatgagga	aatatcttga	catctacaga	caaatatTTg	ttgccaatc	atTTTgggca

22381	ttgaggattc	acaagaataa	ttaaaaaacac	gtaatatataa	atccttctga	ttagagcaca
22441	ttttatggca	ctgcataaag	agatggctgt	gggagtcatt	tatattttaa	tactaacttg
22501	aacaaggctt	ggttctgcaa	ataaaacaga	agagtaaata	tgaagtggag	cacataataa
22561	tgttactgtt	tctgtgggtg	agcttccata	ggaagagggc	aaaggcatca	atctcaaaaa
22621	gtagaaaaaa	cccaaagaaa	gaagaagaaa	aataacaaat	gagaacaaaa	ttagaaaaat
22681	tgtaaataat	atggtagaga	agataaacaa	aaccaaactc	ggttctctga	aaagattagt
22741	tatgtgtaag	tcactgggtg	ttacagagag	aaaagaggga	aggttaaaaa	aaccaaagtc
22801	aacaataaaa	agggggctat	tactacagat	gctgtagaca	ttggaaggat	aatagaatat
22861	tatgatcaat	ttcctatcaa	taaattgtgaa	aaattagttg	aaatacataa	ttggaaaaata
22921	taattttacca	aaactgtcaa	aagaaaaaac	acaaaaacctg	aactcctata	attattaaag
22981	acattgattg	aaaaatctca	gacctcttaa	aggacaactc	caaaaaataa	aaaaataaaa
23041	aaagaaagaa	agttccaggc	cagacggatt	tatcagcaaa	ttctaactatg	catttaaaaa
23101	aaaaaggatg	ccactcttac	ataagctttt	cagataaaag	aaaagaaagc	ttgataccaa
23161	aatgtagcta	gaaaattatt	taaaaaggaa	aattatgggt	caatttcata	cataaacata
23221	gatgcaaaat	cctaaataaa	attctagaaa	acaaaaacca	acagtatata	cgacaataat
23281	atatattgtg	aacaaactgg	atttatttta	agaatgaaag	cttgtttcaa	tcaaataatc
23341	aaattttacat	tgtaaagaaa	atttatataa	ctttaaacaa	aagaagaaaa	gcacttaata
23401	acattcaaca	ttcatattaa	aaattcttag	taaattagaa	ataaagagga	agttttataa
23461	tcctacaaat	tgataaaaag	aatctaccga	aacctacagt	aaaccacaca	ttttcatgtc
23521	aaaactttta	aatttaatca	tttgaagttg	ggggaaaaaga	caaagaagcc	ccacatatta
23581	ccgcatttgt	ctgaaattga	attggcagtc	ttagttagta	aagaaggcaa	ggaaatgaaa
23641	tgcaatgtta	aaggattggg	aaaaagtaaa	ttaaacttta	tttacagatg	acatcatttt
23701	tttacacaga	atttccaaga	gaatttgga	ataagtcatt	agaatttaata	aatgagtttt
23761	aaaagtttgc	tagatacaag	gtcaatatag	aatttttagtt	ttatataatc	gcataaacat
23821	caatagaaat	caaaattttt	aagtgatact	gttttctatta	gcattaaaaa	taccttaca
23881	tatggacata	aatctaacat	aatcgtgca	gaacctctac	accaaaaact	acaaaatatt
23941	attaaaagaa	gaccaaata	aatggataga	tattttcatg	gattgaaaaa	cttaataatc
24001	aaaggtagaa	ttttcttcaa	attgatcaat	agattcaatg	cagatgcatt	caatatttca
24061	acaaagtttt	tgtgaaactt	gagtgattct	aaaatatgta	tggaaatgca	gagtcaaaaa
24121	cagacaagat	gctcttaaag	aagagtgaga	aatacaaaa	attatcagag	actattacaa
24181	actgaaaaac	ctagaggaaa	tgtagataga	tctgaacaca	tacaacctgc	caagattgaa
24241	tcagaaagaa	actgaaaact	taaacagacc	aataacaagt	aataagattg	gataggaaaa
24301	aaaatctccc	aacaaagaaa	agtcaaggac	cagatggttt	cactactaat	tctaccaaac
24361	ttagaaagaa	gaactaatat	caatcctcac	caaggccattt	caaaaaatta	aagaggagag
24421	aattcttctt	aattcattct	atgaggccaa	cattaccctc	ataacaaaa	cagacaagaa
24481	cacaacaaaa	aagaaaacta	caggctgata	ttcctgatga	acatcacttc	agttcagcat
24541	ggatacaggc	caatatattt	ggtgaatata	agttttctag	tattttgttc	agcaaaagct
24601	atcaacaaaa	tattagtagc	ctgaaaccaa	cagcacatca	aaaagataat	gcaccatggt
24661	caagtgggat	gtatcccaga	gatgcaagga	tggttcaaca	tacacaaatc	aataaatatg
24721	atacatcaca	tcaatagaaa	aaaacaaaaa	ccatatgatc	atatcaataa	atgcagaaaa
24781	ggcatctgat	aaaacttaac	attgtttcat	agtaaaagct	ctcaacaaac	taggtacaga
24841	agaaacatac	ctcaatatat	taaagtccat	acatgattaa	cccacagcta	acattatact
24901	gaacggggaa	aggcggaag	tctctcctct	aagaactgga	acaagacaag	gattctcatt
24961	tttatcactt	ctattcaata	taggactgga	agtcttagcc	agagcaatca	ggcaagagaa
25021	agcaataaag	gtcatccaaa	ctggaaaaga	agtcaaattg	tccctctttg	gagatgacat
25081	gatctttacat	ctagaaaaac	ctacagactc	caccaaataa	ctcttaggtt	taattaaaaa
25141	attcagtaaa	gttgcaggat	acaaaaaaat	cagtagcatt	tatataccca	ataatgcact
25201	ggctgagaaa	gaaatcaaga	agaaaatccc	atttacaata	ggtataagga	aaaaatatct
25261	aggaaaaaaa	ttaaccaagg	aagcgaaaaa	tctctacaaa	aaaactacaa	aacactaatg
25321	aaagaaatgg	aaaaggacac	aaacaaatga	aaagacactg	catggtcatg	gatcagaaga
25381	actaatatca	ttaaaatgac	cacacaacct	aaataaatct	acagattcaa	tgcaatccct
25441	gttaaaatat	taacgtcatt	tttcacaaga	atagaaaaaa	acaatcctaa	aatttatatg
25501	aaagcaaaaa	ggagcctgaa	tagctaaagg	aacctgaac	aaaaacaaca	aaactgggag
25561	catcacataa	cctgacttca	aaatatatta	caaggctcta	gtaacaaaaa	gagcatgata
25621	ctagtataaa	aacagacgca	tagaccaatg	gaacagaatt	gataacccag	aaataaatcc
25681	acatatattac	agccaaccga	ctttcacaac	aggcatcagt	aacatacatt	gggaaaaaga
25741	cacctctctc	aataaatggt	gctggaataa	ttggatatcc	atatgcgaaa	gaataaaact
25801	ggacctctat	ctcttgccat	atataaaaa	caactcaagg	tagatgaaag	acttaaaact
25861	aagaccccaa	attataaaac	tattaaaaga	aaatagagaa	aacatttcag	gacattgatc
25921	taagcaaaaga	tttgatggct	aagacgtcaa	aagcacaggc	aagagaaaaa	agacaaatga
25981	gactatatta	aactaaaaca	cttctgcaca	gcaaaggaaa	gaaaacaacc	agcagagtga
26041	agagacaacc	tggtgaatgg	gataaaatat	ttgtaaactg	ttcatctagt	aagagaaaaa
26101	ccagaatata	cacaaaactc	aaacaactca	acaattaaaa	agaaaacttt	tacaaagtgg
26161	gcaaagaacg	tcaatagaca	tttctcaaaa	gaagacacag	gaatggccat	caagcatatg

26221	aaaaaaaaata	ctcaatatca	ctaatacatca	gggaaatgca	aatcaaaaacc	acaatgagat
26281	atcatcatac	cttagttaga	atggcaatta	ttaaaaagac	aaaaaataac	agatgctggc
26341	cagaatgcag	agaaaaggga	actcttctac	actgtgagt	ggaatgtaaa	ttagtacagc
26401	cactatagaa	aacagtagag	atttctcaaa	aaactaaaaa	taagactcaa	tataatccag
26461	atatctcact	actgggtatt	tatctaata	aaaagaaatc	agtatatcaa	aaagatactt
26521	gcacccacat	gtttaccgca	gcactattta	ccatagcaaa	gacatgaaat	caacctaaat
26581	gtccatccac	agatgaaatg	gtaaacaaaa	tgtgggtacat	atacacaatg	gaatactatt
26641	tggccataaa	aataaggaaa	tcatgtcact	tgagcaaca	tggatggaac	cggaggtcat
26701	tatgtcaagt	aaaataagcc	aagcacaaag	gacaaatacc	acatgttctc	actcctatgt
26761	gggagctaaa	aatgttaatg	tcatggagat	agagagtggg	atgatgggtta	ccagaggctg
26821	agcagagagt	gggagtggag	aggaagagag	tttgataat	gggtacaaac	atacagttag
26881	atagaaggaa	taagtctga	tgttgtagat	cagagttagct	gactatactt	aacaaaaatg
26941	tattgtatag	ttcaaatacag	ctagaaaaga	ggacttgcaa	tgctcctgac	acatagaaat
27001	gagaaatacc	caaggtgaca	gacaccccaa	aaccctgact	tgatcattac	acattctatg
27061	catgtaacaa	actatcacat	gtaccccata	actatgtgtg	tgtgtgtatc	tatctatcta
27121	tctatctatc	tatctatcta	tctatctatc	tatctatcta	tttatagtat	ccatttttaa
27181	aaggcaaatc	ttaaaaaaaa	gagtgaggta	ggacaacctg	ctctactaaa	catttttata
27241	aaggcacagc	aattaggtca	gaatagtgtt	catgcaggta	gaaacaaatc	agacaatgga
27301	aaagaaagga	aagttcagaa	acagattcac	acatatatgg	acacacttga	tttttgacga
27361	aggtggcagt	tgattttttt	ttacaaaggg	gcagagttagt	taaaaaaggg	agaaactggt
27421	ggtaatgat	gctggaaaat	gagaaaatcca	catgagaaac	agaaaacttga	gaagtaactc
27481	ataacatcac	aaaaattgat	gtcagataaa	ctataaatct	aaatatgaca	agcaaaaact
27541	atgaagtttc	tggaaacataa	tatagaaaac	tatattcctg	acttgggaag	ggaaagattt
27601	taaaaacaag	gcacagaaac	actaagagta	aaggaaagac	taataaattg	aactgcatta
27661	aaataagaaa	aaaattagga	acttaaat	tcatcaaaag	acactattaa	gagaatcaaa
27721	agggcaatga	tggagtggaa	gatgatattc	ttaacacatc	taatctacta	ttcatatcca
27781	gaatatgtta	agaaccccca	tcaatcaata	agaatacaca	caagtacaga	tacacacaag
27841	gaaaaacagg	gaagaaactg	cacaggcagt	cataaaagag	gatatccaac	tgctcattaa
27901	aaggtcttct	acttcattag	taaagcagga	agtacaaat	aaaacctgaa	tgaaataaca
27961	ccacagatac	atcagaatgt	ctaaaacaaa	gactaacaat	gttggtgaag	atgtggagca
28021	ataagaactc	tcacacacag	ctggagttta	Raaaacatcc	ttatttatcc	atcagggctg
28081	taataagagg	catatagcat	attcaaatag	gtaagtgaac	aaaatttaat	gaaagaattt
28141	acaacggtta	aggtagaatt	tagtaaaaac	caaagggtga	ggaagtcccc	caggctagca
28201	agaacaggga	gccattacta	ccctcaggcc	tacaagggca	agggaggtag	cagtcacgca
28261	gcaccggtag	ctgtgaacac	aggagaggat	aaacaatagg	agctgtggcc	ttccatggag
28321	ggaaaaagtc	actgccaaac	catgacaaaa	gagagaaagc	tggagaaaaa	aactcccaa
28381	gctctctcct	cctactctca	ggcttccttg	tggtgactct	catttgctta	acccaacaga
28441	agatagagga	caaaggagcc	catttgatgt	agtcgacaga	agtcaaaactc	ccaggacaag
28501	aggagcagag	aaggatggag	aatgaattag	agaagttaac	aaagaaagtc	ggcacatcta
28561	ctgtcatcaa	agactgtgta	attccactat	ttaatacatg	tacaatagag	atgtgtgtat
28621	atgtgcacca	gaagtacact	acaagaatgt	ccactgcagt	attattctct	cagttaaatg
28681	ctatatctgg	ctgcacaact	agattaggac	tgagacccta	accacacatg	tagaagcagg
28741	gagaaagatg	cagttgagaa	gccctttgcc	atctgatatg	cttccaaaca	agggaaactc
28801	agcttgagca	ctgaaagcat	tttgtataaa	atactgtttg	tatgatatcc	ccatgtcacc
28861	acacacacac	acatacagac	acacacacac	acacacacac	acaacaaaaa	caaaaaaaac
28921	ttccagggtt	acttttctcc	tcagttagca	ggtcaaaact	tttcacgtgt	gaaatacttt
28981	tgcaaggagg	cagatgaaac	gggggagaa	ggataaacct	gccttgcatg	catctatcaa
29041	ttcactgcaa	ggaaggaaca	cagagcttag	atatgaagca	tcaggagttc	ttcccaagac
29101	ccttcttgtc	ccttgggaag	agccaggagc	tggaggggat	ggccttctact	gagaagtggc
29161	atcaaattag	aacctgcaca	acaaatagga	accagctact	taaagatcca	gagaagttagc
29221	cctctagact	gaaggatcag	ctcctcaaat	atcctttttg	attttgaaact	cacaagttca
29281	ggggcatccg	aggaacagac	acaagccagg	gctgaacaa	ctcatctgtg	tccatgaggt
29341	gggcaagagc	cagatcctac	tgaggagtgt	gaacttaatc	ccaaacatgg	cgggaagcca
29401	ttggggccaa	agatgaatta	gggattgtct	gatctccttg	acacgttctg	aaggttttgt
29461	gtcacRaana	ggtttgtcaa	catgcacaaa	ggaaagctgg	gataaccaag	ctaagacgct
29521	accgcatcaY	tctaggggac	gggtgactca	ttccaggagt	ttctgcagga	agacatcata
29581	ggactggtga	ttgactggag	gcagagagca	cgaaaaaaga	acaatcagag	atgatctgtg
29641	gtttgtggtt	gttgtgtaat	gttaaggctt	aacaactgcg	aaaatagtgt	ttttacttta
29701	cagagaagaa	gacttgaaga	atacaagatt	ggggagagca	gatcgggatt	tgtactttat
29761	atcagacgtg	ctcatagata	tacattgtga	gatatcagtt	gggagctgga	tatatcagga
29821	ttcgagactg	ggaatcacca	gtttatagat	ggtttttaag	ccatgtgatt	ggatctcatt
29881	tcttagtggg	gagaagttag	ataacaaata	aaagccaaga	ttaaagctga	accttcagcc
29941	tgcccatcac	ttagaagccc	agcagggaca	gaagagccag	ccgacaRcac	agagtgcagg
30001	aggaggattt	cYgggcttca	gtgagtgtgc	atcagagatt	catggacaca	acttaaaagta

30061	agaccagcca	cacagttccc	ttcattttct	caagcaacat	tcagcatctc	aaagtgcggc
30121	tactacacac	ggagccaagc	cctgggtccag	ttatcacaaa	accttactag	atatgtcgat
30181	tatttagctgc	attttatagc	taaaggtagg	caggcacaga	gaagttacgt	ctcttgccca
30241	cagtcacagc	aagcaaacag	acctcagacc	aaagtctttc	taactcctta	tttcacatct
30301	aaaacactat	gtgacgccac	cKccaaaaaa	tgtgggggtt	actctaggac	aaaagaagct
30361	catgaaagca	aagagagaag	catagagctt	aaatttagaa	caatattcag	ggtgattggg
30421	aaaaatagcc	tccatacaaa	tataatcggg	tcataccag	gtgaataaaa	tgagagagga
30481	acctggtaca	agtgaagttc	attgaatttg	cctgcaaaaa	aacataattc	aagtgtggaa
30541	gtttggaaac	taaaatagta	atcagctaaa	gaaattttca	ggccaactta	gacacaatgg
30601	tgacaaagct	cccctgagga	gacaggaact	atactacttg	caactctaac	tgaatggact
30661	taacttagca	attaagatca	ccgtaagttt	gatcagataa	gtattaaagg	gtgtcttggg
30721	tacatttgca	gattcagaca	ctcttagttg	tgaaattctt	caccaaaaaa	ttgcaactta
30781	tggcttgcac	attaaaaaca	tgatatgaat	gaaattcaga	tatctcttgg	taaattttat
30841	aattattcct	cccctcagct	aaagttctca	gtatgaaggg	ttataactga	tgtcatatac
30901	aataactaca	agctgtatca	tacaaaaatt	gagaattttc	tattatagaa	agagttgctt
30961	taatctaaac	ctaaagtagc	atcaataatt	tttcgtaagt	gggtcccctt	ctatgagtaa
31021	gccccagga	aaaataaatt	gcttttaaga	aatgctgtaa	gtaagctggt	tcatatataa
31081	gttttgacaa	gatctggcta	cttaaaattac	aagatataaa	atgagatttg	tactttttct
31141	caaatgtaaa	aaaaaagatt	tactttctgaa	ataatggtct	ttataaaata	tgcatttttg
31201	ttcctttttc	ttatcatatc	tataaattta	ctaataaaag	attctatctt	gattttacta
31261	catttgaaaa	gattgattat	attgggtattg	caagaagagg	ataccataaa	ttaaaaatcc
31321	atatattaac	catattcttt	tatttgcaat	atgacaaatg	atttgttaaa	gttataatca
31381	ataaaagcta	tttgaaagtc	ataaaatgga	tattgccata	actgctttcc	tagcaaggtt
31441	aaaaacattt	aactattttt	aatcaactta	cagtgtagta	gatcaagcat	atctaaatat
31501	tttattttta	tcaagcaaat	atatttctgc	tgttcctatg	acagtatgaa	agtttggttc
31561	aaaatatttc	caagggtattt	aacacagttc	agtagaggtc	ctgtctcata	aatcaatttc
31621	agattagata	aaaggaaaaa	gcaaatctct	gtttacttca	gaatttttca	agccagttcc
31681	aaagacagag	ttgttactct	catgtgaatt	ttcagtagaa	agctgaacaa	cgacgttgct
31741	ctgacaatct	ttttaaatat	catatgattg	agacagaata	acaaaaacaa	agtaggtata
31801	agctgtacga	aattatttct	cttggacatg	aaatatggtt	ggggaagaaa	ggtgtgttct
31861	catatggact	agggtagtta	cttctcaaa	tgaggagcaac	tcagactctg	aaaaatggtt
31921	tctcaaagga	ggcatttgtg	ctgggtgagg	ccagttttct	agtcaaaatc	attttattaa
31981	atgccattgg	attgaatgct	tccaggtttg	ctctctgaaa	tatgccctat	cattctttaa
32041	taaacacact	aggtagcaaa	aagggcacac	taaaaataac	gtgggttaat	gtttctatcc
32101	ggcctgtggt	ttctaagtag	ccctgtcctg	agctttcgac	ttttttaa	gaacctcaa
32161	atagttggtt	gtctatcaca	aataagccat	ctattttggt	attcctgtaa	cttttYgtag
32221	aatccagaaa	gtaacattct	tgctgatata	tcttatataa	caccaagtc	attttccaaa
32281	tgaggattcc	aatgagtatt	catcaccac	atctccccga	aaaatattta	gaaaaaacag
32341	actcaatcct	ttatccaaag	tcagttagta	agaggcgggg	gccaggggtg	gacaaaaatc
32401	tctttggcca	ggaagccggt	ggtcaccctg	atctacttct	ctaagctccc	acccctgatg
32461	aatgttcaat	atctgagttt	cacccactc	attattgtaa	ttctccgtaa	aacccYagtt
32521	accacacagg	aattctcatc	tacRtctagc	ccaatttgct	actggatttc	aatgttattt
32581	catatgaaaa	caatcactaa	cagaaagaac	ctggatctgg	tttcgaaaaa	ataatttgaa
32641	aggaattcct	tctgaaggct	gcaaaagccc	agtgatggcc	catcggtacc	ctgaatgggc
32701	ttctcagggt	cctccactct	cctcgtatg	tctccatcca	tctcaRaata	catgtctggc
32761	tcttctgctc	cataaacaga	aacagagagt	tccgggatcc	aaaggcaaga	ctgcgattta
32821	agagacactt	tgtttttgta	ggtgttcaag	gatttggaat	caaagcagca	gaattattaa
32881	aattaatcca	attttcttgc	atttattttg	ttagttcata	cctctaaaag	atatattttg
32941	gaaggcaccc	ttcaaccac	ccattttaca	atcaaagctt	tctttggaat	ccctgagacc
33001	aaaccaagc	tagacagtga	cacactgcca	gctccagagc	cctgcccacc	tggaagggtt
33061	atggaatgca	ctatttgact	tgagctcaca	tttaattaca	gtgtcaaaaa	tcaaacttaa
33121	attgctttta	tcagctKcca	acactatgag	cctttctggc	ttgaatttct	ataaacaaag
33181	aataaattca	aatctgtcac	aaagtggcag	caacttggtg	accactggac	taggaatcgt
33241	tatgcaacca	tggctcactg	tggaaaaagt	caacatcttc	ctatatgtga	ctcttttctg
33301	atcttttcgg	actattagtc	aacagaatca	tcataaagct	aaattatcta	ctgcactga
33361	aattactcag	cctcagacat	ccaaaccagt	agctgttcac	caatgaaacc	tgaataaagc
33421	tgttgctgac	tttccctgct	tcaatttaga	gtatttttgg	aaacatcatt	gcatgtttaa
33481	aatataataa	gcaaaaaggac	cttttggtaat	taaacattct	gtaattttaa	actccaaRag
33541	acaggagcag	tgataagata	tttagattttc	tttactgggg	tgaaaaagac	agaaaaggtg
33601	taactgttta	aaaaaagaaa	gaaagaaaaga	aatccaagga	tttttagagc	cagaaaacagt
33661	aacagtgtcc	tttattctgg	aaatgtttta	aagcctgtaa	gcagagatca	cccagttcat
33721	cctgagcatt	ggtgagcaaa	taacttctga	aatgctcttg	aaagagaatt	ctccagcctt
33781	ctctgtttga	aatacttYct	aMagcaaaaca	caaaagtggtg	caaaacacaa	gagctcatgt
33841	taattccaaa	caaagagacc	aacagtcgca	aattgagagt	ccatcttaga	agcgggtgct

33901	gcttccacat	caagcaacta	ttggaaaga	taggggtaga	tgcagaatta	agatcatttg
33961	tgaYtacaga	aagagacaag	aaaattaatt	tttaaaacta	aaagagaaat	ttatttcaga
34021	taatccaagg	aaatgagcag	agtcaaccta	acaggattct	caagcccctc	cctactggtg
34081	tggtgtgaag	atcacacgtg	gtgtagctga	caccctcac	atcttcccgc	cRccctcct
34141	gccacctctt	tttggcattc	accctctctc	gggttctgta	tccctgccca	tgttctccat
34201	gaaaggccaY	gttaatgccc	atgctcttct	ctgcctctaa	cagacgatgc	cttccacagt
34261	cactcaagct	tttaaccctct	tctggattcg	gacctgtaca	tccaaaagcc	tgccagRcat
34321	ctcttccaga	tgtcttcaga	gcacatgacc	agaagcacgc	gcaaagctct	cctgagcctc
34381	agtgtcttcc	ctccaacact	tcttccactc	tctcagttac	cgagccaaga	cacagagttg
34441	cttgtgagtc	atttctctac	ctcatcccc	acccaacctt	cctgcgagat	cctgctcaYg
34501	ttacctccta	aatgctgtta	cgtgacttca	accacgtcag	ctttttatgg	gggaggagag
34561	gcacctgcat	atctgtggga	gcaccaacct	ctagacatct	cctgccaaatg	tactgcaatg
34621	tgcagaaaag	aggaagaaag	caagcccaag	tcgcaaagtc	tttgccagtc	cttctacca
34681	ccacaccttc	ggtggagatg	tgggtctgtg	gaaatgcaga	gacggactga	aagccaaatg
34741	ggccaatgag	agMtctgaac	aaaacattaa	gcttgttttt	ctttaaggca	atagcattaa
34801	cagtgggtcc	tttaagaatcc	gtttaaaaaa	aataataata	atacYttttc	tgactattgt
34861	tggggaagaa	taactccttt	atggggagta	tgaactcaaa	gacatgatat	tgtaaaggag
34921	gccttgaaga	agagagggac	taagaaatgt	gatggcgtgg	tctgttacat	gtgcaggtgc
34981	tgtgtgtgag	ggaagcactg	aggaccactg	gaggatgagg	gatgctctgt	ccttctcggg
35041	tatggctgaa	ggacagggat	tgcccatcaa	ggagtaaata	atccatgagt	ggactaaatt
35101	ctttagaaac	aggaactcaa	atccagtaga	gattaggata	aaaatgtaaa	ctctaatttg
35161	ccctctcccc	tctctcccc	aagtagtcgg	tttctataat	tttataactaa	atcaatttaa
35221	tgtatctgct	ataatattaa	tcataatttt	gtctgaagta	gtaacaactg	gaaactacct
35281	aaaagagtca	gagcaggact	gtaagactat	attgtagtct	acactatgtc	acagYggact
35341	gcatcatgat	agagaacgat	actcatgatt	taactgataa	acacatatgg	caatgaataa
35401	aactgtgagt	aagatgtctt	taattttagc	aaaattatct	tcattctaata	tttcaagtaa
35461	acttaagatg	ttatatataa	atgcacataa	tttaattaaa	atacatgcat	atattactaa
35521	gaaggcta	cccccaaat	gtccatagat	gttatttctg	agtgaggagg	gtcaggaagg
35581	ttttaattttt	ctttgtacta	actctatfta	ctaaaatgtc	cacctggatt	aacgttataa
35641	taaaaatggt	accgttttaa	ataaaatgaa	taaaaatacac	cggtatgaag	gtattattag
35701	aacatacgaa	atgacaaaat	caagcagcaa	aaggaggaaa	ttatattaa	aactgaaaaa
35761	actatgttga	gaattctgtc	taagaaacgc	tgtagcaatt	aagaaaaatc	ccagtctctg
35821	aagtaatggg	agaaagatga	tgacttccag	aactttcacc	atctactaag	atgttcacca
35881	attcctaaga	agggactgaa	tacgtctgca	tcctcagatc	tgagagtaat	tttccacgac
35941	gcatactgag	caccataatg	agtgaagtct	ctgagatgac	ctctctgacc	tctctaaaaa
36001	aatcaggcac	tgcaaacatg	gcctatccca	cactagccct	tggtgaaatt	caagggcata
36061	atattagact	gaaacagtga	gaagaagatt	cgacatggag	tgcggaaga	atgtgttcta
36121	gtccttgaag	gccagcgcta	ttccaagtgt	ggtctcagac	cagcagcctc	agctgcagct
36181	gggggcttgg	aaatggaaac	gtatgtgcta	cgtcctggac	ctgtggaatc	agaatctctg
36241	aggtggacca	ggaatctgtg	tttttaaaac	cctgtgggtg	gtaacttctc	gctggctgag
36301	gacagagaga	atgcttagtg	aacgctgagg	aatgacgggt	taataaatcc	ctgaatgatt
36361	cctgagcgga	tcaattttctc	agccagagtt	ttgaccgcag	ctgatgatga	acagggctgt
36421	gaYcaccagg	aagggtctga	agaggaggcR	ttgattgcag	ttttctgcct	ctgatcctca
36481	Kgcatggagc	cccaccccc	aactctctta	cacggaggcc	aaagaaaaag	ctggaattct
36541	ccaggagagc	aggacctgtt	aaacagcaaa	gggcaaccca	aaattgtatg	agtggccttg
36601	ccctgagaag	ctttgaagcc	agaggttact	gaacaactac	cctggttaaa	accctaacgt
36661	gacaatgata	acatgtgact	gtatgtggct	gtcttcaaaa	cagcaatgat	ttaggcacaa
36721	ctgtggatct	cagattgttc	cgtgagtttg	gtctgataag	attgtatgcg	tttcccaggg
36781	gtgccataac	aaagtgccac	accctgggcg	gcttccgcca	ccgaaatgta	ttttctcaca
36841	gctctgaggt	cgaaagtcca	agatcaaggt	gtcagcaggg	ctggttcctt	ttgaggctgt
36901	gagggaggat	ctgttctagg	ctctctctcc	cccagcttct	gggggctgct	gcagtctttg
36961	gtgttctctg	gcttatagaa	gcacatcccc	aatctctgct	tttgctgtca	catggccttc
37021	tctcatata	cctgtctgtc	tccaaatttc	ctcttctcat	aaggacacca	agcactggat
37081	taggggcccc	ccctactaca	tctgcaataa	ccctatttcc	aaataagggtc	acatttttag
37141	ggaccagcta	aatacaaaac	attagccagg	catgggtggg	ggcacctgta	gttccagcta
37201	ctcaggagaa	tcacttgaac	ctggattaga	acttcacaat	tactgacacg	attcaaccca
37261	tatttaggttg	tttagtatta	tacattagga	agccatagca	aacttttttt	tttttttttt
37321	tttttttttt	tttttgagac	ggagtctcgc	tctgtcgccc	aggctggagt	gcagtggcgg
37381	gatctcggct	cactgcaagc	tccgcctccc	gggttcaacgc	cattctcctg	cctcagcctc
37441	ccaagtagct	gggactacag	gcaccccgcca	ctacgcccgg	ctaatttttt	tgtattttta
37501	gtagagacgg	gttttcacct	gtgttagccag	gatgggtctcg	atctcctgac	ctcatgatcc
37561	accgcctcgg	gcctcccaaa	gtgctgggat	tacaggcgtg	agccaccggc	cccgcccat
37621	agcaaacttt	tgtaaacata	actattttatt	aaaatttggg	ataattaata	tttaaatatc
37681	ctttctggca	agatttagcc	agttttttta	gcttcattga	gagaaccaca	atattagcat

37741	taataataga	atgggtctatt	aatttttttaa	aaagatacta	tgtctgcatt	tccaatataa
37801	aattcacagc	aatgattcag	gtagaaatta	aatgaggatg	tcattgtcaa	cctttttaa
37861	tatattacag	gagcaggag	cacacgggcc	ctgaggtaga	aatgaataag	gaatattcta
37921	gaaccaccaa	gaagccagtg	tggccacagc	agagttagcc	agtgggtgctg	gaggcagggtg
37981	agatcttgaa	gggccttggc	aacatgaaca	ggagcttgga	ttttatccac	attgcccagg
38041	gaagtatgag	gtgggtgttta	gcagggcata	aatctcatct	gatctgtact	ttgaggaaaa
38101	cactctgact	tctgtgtgga	gaacagtcta	taggaagcca	agaggagagg	cagggaatg
38161	agggaggggt	gacttcagtc	tccaggagaa	tgagctagt	atgtgaatta	aggtgctggt
38221	ggtggaggaa	gtgagggtgga	gaactggaca	ctgggatcta	ttttgaagag	ggagcctaca
38281	gaactgtgtg	gtggatggct	atgagggaaa	gaggattaaa	gaatgatgct	aggcatttgg
38341	gctttaaaca	acccttgagt	aaatggttat	cctattatta	atactaagat	gggtaatgct
38401	aggtgaaggag	tagccttggg	gagcagtgat	gagttatttt	cttagaSaga	agccacgtta
38461	cctataggag	gctcaacaga	gtatcaaatt	agaataaaat	atctaattgtt	ctgatttgtg
38521	tctttatttt	ttccattaag	atcacatgg	ctggccaggc	gcggtggctc	atgcctgtaa
38581	ttccagcact	ttgggaggct	gaggcaggcc	gatcacctga	ggtcaggagt	ttgagaccag
38641	cctggccaac	gtgggtgttac	cccgctctgta	ttaaaaatac	aaaatatag	ccaggcatgg
38701	tggtgggcac	ctgtagtctc	agctactcag	gaggctgagg	taggagaatc	acttgaaccc
38761	aggaggcaga	agttgcagtg	aacYgagatc	acgccattgc	actccagcct	gggtgacaag
38821	agcgaaactc	cgtctcaaaa	ataaaaaataa	acaaaaacca	tgggaaagaa	caattgagga
38881	aactggggac	ttgtcctgga	acagaagaca	cgaggacttt	tggtttttta	gttttgtttt
38941	tggttttgag	acagagtctt	gctctgttgc	ctgagctgga	gtgcaatggt	gcgaactcag
39001	ctcactgcaa	cctccgcctc	ctgtattcaa	gtgattctcc	cacctttgcc	tcctgagtag
39061	ctgagactat	aggcgcacac	caccatgccc	agctaatttt	tgcattttta	gtagagagag
39121	ggtttcatcg	tggtggccag	gctgtcctca	aactcctgac	ctcaagtgat	ccaccgcct
39181	cggcctccca	aagtgcggg	attgcagaag	tgagccactg	cgtccagcaa	gaatgagggt
39241	tgaaggagg	catgattgtg	ttctcttcaa	atagttgaaa	ggttccattt	gaaaaagga
39301	tttacttatt	cctatggttc	tagtaagttg	cagagggtgg	atctaagacc	attgggtgaa
39361	ctttacagag	aggccagaaa	gccttcWcca	ggggaagaaa	gtcaagaaaa	ggtctcaaaa
39421	ctgtctggca	gaattctgca	caggagagtc	ctcaattYgg	tgtagaattg	gcaagtataa
39481	taactgtaaa	ttccctctcaa	tctaccatct	acatcattag	gcgacatctt	catttcctga
39541	aaattccaac	acatgaaggc	tcacttacta	tcaagatgag	caagtttgtg	ctacatagt
39601	aatttgattt	caataggact	cattttatac	cttttaagag	tctttaaaac	gttgaaaatg
39661	taaagatctg	caagagcaac	tgatcttcaa	agtcccgtcg	agtcacagga	ttttgtaatg
39721	atgtctaaag	agggaaaattc	ccatttcaga	aaaggctcaa	tagtatttaa	tagttagctc
39781	aaaaggacga	ataaagttct	acatgggaag	ctgctgaact	aSagtcaggc	caaacccttca
39841	tactcggaga	agatgactgg	gctgtgagac	acagatattg	atgataaatc	atcaacaacc
39901	acagcttttc	ttaatgggg	tggtgggcag	ggagagcagg	cttcagggtg	gccttcctcc
39961	agcaagggca	cacacacaaa	ggggacctga	acacgtggag	cttatgggtca	ggaagcctga
40021	agtgcctgac	ccagcacatc	tggtgactag	tagtgtagat	tccccacaga	aggaatgaca
40081	aggcgtcatt	attcaaagt	accttccctc	cattggaata	atcagaaaca	gctcacaaat
40141	aaagtaatct	gtaagatgct	tcacttcttc	taggtaacct	ctccccaa	tgtatgacatc
40201	acttcacatg	aggcgacatg	gttttgtgag	gacttttctt	tccttccaag	tttggaatct
40261	ggctgaaggc	taaaatcacc	atttacagt	tttgagccct	ccttSacact	ttgaagaagg
40321	gtctgtgtaa	atagccattc	ctacttctca	gactcgccag	aggcaaaacc	tgttttgctt
40381	tggtagtaca	cgctcttctc	taactggtaa	aagactttta	aaaaaaataa	gatgccattt
40441	ttaatgtttt	cataaatctg	gcttagatat	atlaaggaaa	ttgctatttt	atatgtagaa
40501	cagcggaag	caaattctcaa	gagattaaaa	atcatgttct	tgtgggtttc	ttttagaagc
40561	tgcaaatttg	aattcaattt	aaaaaatcat	ggcacaggga	caaaataaga	atgaacaaag
40621	atgggaaaaa	tgtccgcagc	attctgcaga	gtcaagcgct	gctgtgagtg	agtgtgcatg
40681	tttccaatta	cacatcaggg	cttggcgagc	tccaggagtc	aaattcagac	agaacccaag
40741	tgcagtgcgt	tacacagaac	ccaggccagg	aaaagcttca	cagagtcctt	ctttcaactg
40801	aggcaattgt	gtagacaatt	tggttcattt	caagaaaaaa	ataaaaactga	cagaagaaag
40861	aaaaatatYa	ttatatattt	cctctcaaaa	caaagtcaaa	ttgtgttggt	gtataaatca
40921	agattYagta	tttccctggc	tgcttccaga	ataatagtag	atcacaaaga	gaataagaac
40981	agaaagttc	agttctcagc	cagacgatcY	aaggtaacaa	agggtgagac	cctcatttccg
41041	ttactcacag	gagcaaacag	acaaccaata	ataattgggc	acatgggggt	ttcaacaaga
41101	taaactgtta	tgtgagtcca	gggtgcta	ttcYggattc	ttttgagttg	caaggaaaa
41161	atataccgat	ttcccccccc	cacccaccca	cataacaaga	cacaatagtt	aacagctgga
41221	agactgtggt	tagggggaaa	tgtgattgta	ggcttccatt	tcaaatgatg	gaaaatcatt
41281	tctgaagtca	ctacgaaagt	tacttcacaa	aactaaaaga	agccttgaca	ccttgtgtta
41341	ttccatcatc	ttaaagcaatt	ctagtttggg	aacacttgat	tcttggcagg	ggatttctaa
41401	gaagggtgccg	tatgtgaaag	gctggtgtgt	cagaaggagc	tttaagggaac	tgggaatgaga
41461	acagccaaaa	gctaagattt	ctgattataa	tcacaggtct	ggatgtgttc	aaggtaaaaa
41521	gagatcatcc	ccatcatttg	caatatatttc	aagagaatgg	ggcttttagtg	acactgatca

41581	aagtctgaca	tggatttttca	gctgccttcc	aactccacat	acctgcaatt	caatgatctc
41641	tatggatagt	agaagtaact	gccaccctga	tgttctcacc	ggatgatttat	acgtgaagtg
41701	gactcaaaag	aaatacaacg	aaatccctta	ataaaaaatta	tgtccatttga	cccaacaagc
41761	ttcccagtg	tgtatcagtt	ctcaaccgat	tttagagaca	tttgggtcaaa	aatgatctac
41821	ttgtaataag	ccatctgagg	gcactagctc	tttttcacct	ctgtgttccct	tgaagcagga
41881	agtcagtaaa	atcactattg	aattaaaatg	aagtaaagca	ctttgagatt	tcagatcttct
41941	aaattaagga	acacacaaca	tacgggtactt	tatttgcaca	agaRttatta	gaacccagg
42001	ctaaccacca	aatgaggtca	caggtaaagt	aaaatctgtc	tcagttattt	gaactcaggt
42061	caacacacag	acagagcaaa	gcaccagctg	acctctgatg	cacaaccag	tgttcttacc
42121	atagaattcc	agaaggcatc	tcagtcattc	tcttatgtaa	atacaaaagac	gWcaactgg
42181	ttaaaaattc	atattttttt	attttagctt	ctttgtgtca	aactcaaaga	agcctcacat
42241	aatgatcaa	agactttttg	gtggatcttc	caaacgtaca	tgccgggaag	gocctgttgc
42301	actttttcct	tcaatcctct	aaaatgaagg	ttttgaagat	gcttgccaat	ggaacgttgt
42361	gccttacagc	tataatattt	tttaaaaggc	attcagttta	aattaatata	tttggccaga
42421	caactgtatt	aaaattcttg	ttgttttttc	tgttataatg	gtagcctgtg	gtatgactag
42481	tcttcttcaa	aaatagtac	agttattgtc	tctctcttgc	ctggggata	tcagtgttct
42541	ttcatttgag	tgtacattgc	aatacctcat	taaaaaccaac	acttaggcta	cacaaaagtc
42601	ttcatatcat	gccgcaaagg	gcgtaaacta	acgtggaaag	tttcagagcc	tcataactca
42661	ttgcatatac	ctcttatact	agctgctcaa	ctgatattta	gctctcctgg	atgatgtcac
42721	agttttttatc	ctcttaaaat	agcctgcctt	caKaaaaat	atagtttccct	tcagtactca
42781	aatctggga	gttacgatgt	aaatagagcc	attttttttc	ttatacaagg	aacaagtgtc
42841	ctcaccttcc	tcccttctct	ccctccccat	acacagaagc	acacttcaact	tcacttatac
42901	taaacacttc	acttatactt	acactgcaac	aaagagtgtc	aaatacatgc	aaataaagcc
42961	tttataactc	ccctctctcc	tgctggacag	ctccttgact	caagtatttc	Ygaagacaga
43021	cactctgata	agggctctgt	ctgagcttca	ttctcagcat	acagtcctgg	gattattttac
43081	ctaccaatat	tcacctctgc	atgaaccagt	aagtcatctac	tgacaagaaa	aaaaacaattt
43141	gtgatttttt	ttaatgtcac	aatttttctt	tttaaRgttt	gggagagaaa	ggtgttttga
43201	atgtttctac	agtttctatt	gccaaatgaa	cacctgtttc	agagtccttg	aaagtcacac
43261	atcagaagct	tctgacttca	acatatagtc	tgctttgaag	aattatttta	aaaaagcatK
43321	gactgcttga	agataagcaa	gtgacactca	atagcttcat	tcttttagtct	tagaaaataa
43381	Wttttacact	tcacacctat	taggatggct	actatgaaat	ataaaaaaaa	ttaaaaaata
43441	caatgggtggc	aaaaatatga	agaaattaga	acccttgtgc	actgctgtgtg	ggaatatcaa
43501	atgatacagc	ctctatggaa	aacagtatgg	tggctcctca	aaaaatttaa	aatagaatta
43561	ccatatgatc	cagcaattct	acttctgggt	atatattcaa	aaaaatttaa	agcaggatct
43621	caaagagttc	tttgtacagc	catgttcatg	gcagcattac	tcacaatagc	caaaaggtca
43681	aagcaaccga	atcgtccatc	cacagatgaa	tgaataaaca	aaatgtggca	tacaataaaa
43741	tgtcattcag	tctttaaaaa	gaagaaaatt	tggacacatt	gtacaatata	gatgaatctt
43801	tttttaaaatt	aaagtgggtt	ttttaatcgt	tttattttta	agttccgagg	gacatgtgca
43861	ggatgtgcag	gtgtgttaca	taggttaaag	tgtgccaatg	tgggttgcctg	cacctatcaa
43921	cccatcacct	aggtattaag	cccagtatgt	attagctatt	ttccctaagg	ctctcccacc
43981	cccaccatcc	cccaacatac	accatagatg	aatcttaagg	acattatgct	gcgtgaaata
44041	aaccagtcac	aaaaagacaa	acagtgtatg	attctactta	tgtgaggtac	ctagagtagt
44101	ccaagtccag	aaagttagta	taatggttgc	caaaagggtg	gRgcattgcag	aacacgggga
44161	gtgttgattg	acatcaagtt	tctgttttgc	aaggtaaaga	agttccagag	atttgctgca
44221	ccaacaatgt	gaatatactt	aacactactg	aactgtaaag	tgggtgtgac	tacagatttt
44281	atgtttttatg	tgtttgcccac	aaccaaaaat	aaagatttaa	ttaaaaaata	aaagatttta
44341	tagcacctat	cacttaccac	acccactgta	catcttacac	tcRtgaatat	tttatgattt
44401	tcaaaacacc	tttatttttt	tcccttattt	gttcattcct	tcagtataaa	cccacaatgt
44461	gttaggttat	gtgctgggca	gcttgggaag	cccaaagacg	aattagggaa	gtaccctata
44521	tttaaagcct	tgcaaggaca	taaataatcag	taataataag	taaaaatagt	tcactgtctga
44581	aaactgctat	acttcctaga	tttttcaatt	caattgcaat	ttctaataat	ttgccatgtt
44641	gatccacaa	attcacttgt	gtttaaaagt	taactgtgctt	gattatcata	cagctgcttc
44701	tttcataaga	ggaatacaga	taaaaYgcta	tgggatttca	gggggagaaa	gaacttctta
44761	ttaaattcat	attttcaacc	aataagaaaa	cacaattttt	tttaataaac	actatcttct
44821	tggtgtcttt	ttcacatctc	atcctctttc	ctcctttacc	ctgcctaagc	aaccttttct
44881	ggaccatc	ttgggttatt	acatctgacg	gcaagtgtta	tgtgggcca	ggaccatgag
44941	caattagccc	cctctctact	cctagggatg	gagtacatgt	gatgatttcc	atcagactca
45001	gattaaatat	cttgctggta	agttaccctt	aaaaatcgat	taatggactc	tccaggtaga
45061	cactgaggt	atctccaacc	aaacctRact	tatacttagg	ggctttggaa	gcccaccaga
45121	gtcatcaagc	catatctcca	Yctaggtctt	atactggaaa	gctgggtcagt	agtaacacaa
45181	agaacaagcc	atttgggaaa	acagcccaca	tgaagaagaac	tgtgggttca	agaagagtc
45241	ttccatccca	ggtttaaaat	actcttctcc	tctttatcta	ccttgaggac	tcctatctat
45301	ccttccatac	cagatccagc	ctcatcttct	aagtgaagcc	ctcctctggg	acgtcagcSt
45361	ctcctctgca	tcccctagca	ttttatgcat	cctctgttaa	agcaactgtt	tgcacagctg

45421	Yatccaccac	taaacttaga	ggcaacagac	catggYtttg	aatgtgYaga	acccaacaca
45481	gagccYgaga	cctagtgggt	attcagtgca	tggtgaaggy	aagaatgaat	gcatgagcaa
45541	acaagacaat	ggaagtctgt	catcagttatt	tacattttcaa	aatgcaaagt	gaaaattgtg
45601	caaatctcta	atcaagctgt	actttaaaat	gctaaattga	tttggctaaa	aatattttcaa
45661	cactgaaatt	tctttttgca	agttaagtct	tcgcataacc	aaaacagaRa	agaacRtatc
45721	aatgaactaa	atagataaat	ttatggctga	caaaatctca	agattttagt	tgtgagagga
45781	aagtaaaatt	caatgtgtcc	ttgaaagtcc	tggggtctaa	gatatacagg	ccaatggaaat
45841	agaatagaga	cttcagaaat	aaactattgc	atataatggtc	tgataacttt	tgaaagtggg
45901	tcaaagatca	tttaatgaga	aaaaaaggag	tcttttcaac	aactgggtgt	gggaaaactg
45961	ggtatacata	tgcaaaagaa	aaaagtggg	ctcttacctt	atactatata	caaaaattaa
46021	ctcaaaaatg	aacaaagacc	aaaaataaaa	agctaaaact	ctaaaattct	tatgagaaaa
46081	ctctaaaatg	cttagggaat	aacttcataa	cactggattt	cacaagcctg	tcttgaatat
46141	gacaccaaaa	gcacaggcaa	cataaataaaa	agagacaaat	ttgcctgcac	taaaattaa
46201	cactttttgtg	agtcaaacaa	gctatcaaca	aagtcaaaag	acatcctata	aaataagaga
46261	aaatatctga	aaataattta	tctaataagg	gattaatatc	cagaaaaaaa	aatatatata
46321	tatatgtata	taacctacaa	ctcaacaaca	acaacaaaaa	caaccaattt	taaaaattggg
46381	caaaggactt	gaacagaatt	ttttccaaag	ataacactaa	tgcccaataa	gtacatgaaa
46441	ttatgctcaa	catcactaat	cattgggaaa	atgcaactca	aatccacaat	gaaatacttc
46501	gtagccatta	gtatgactat	tacaaaaaaa	tcagaaaatt	acaagtgtctg	gtggagatat
46561	ggaaaagttg	caaatcttat	gcactgggtg	tgggagagta	aaatgggaaa	gtcactataa
46621	aaaacagtct	agtgattcaa	aaattaaaaa	tagaattaac	atatgattca	ccaattccgc
46681	atgaagatag	caggaaattg	aactgatatt	tgacatatcca	tggtcatagc	agcatttatt
46741	acaataacca	aaaggtggag	gcaacccaag	tgtccatcca	cagatgagtg	gataaaccaa
46801	atgtgatata	cacatatcat	gggatattctt	cagccttaga	aaaggagaga	aattctgaca
46861	tatgctataa	cacagatgga	cctttggaaa	gttaagagat	tgtcttattt	atcttacgct
46921	aagagaaata	agccatccac	aaaaggacaa	atgctgtacg	attccacttg	tgtgagttac
46981	ctagggcagt	caaattcaca	gaaacagaaa	gtaaaatgat	ggtttccagg	ggctgagaga
47041	agggtagaat	ggaaagctat	tgtttaaatgg	gttgagtggt	cagtgtauga	aaatgaaaaa
47101	agttctggag	atggatgatg	gtgagttttg	caccacaatg	tgaatgtatt	taataccact
47161	caactgtaca	ctaaaaaatg	gttaggatga	taaattttat	gtatatttta	ctataacaaa
47221	aagtcataaa	aatcaaacat	ggttcttctc	attaaaaaaa	aaagaagatt	ctggggtcta
47281	gttggttctc	cctccttctt	agttctttga	tcaaggagat	agaattgtgt	ccaaagtgc
47341	caagtgcact	acggagtaat	gggcttccat	ctagccactc	cctgctttat	gactgagctg
47401	cactgggaga	ggcattgtca	ggggaaattc	caggtttcta	gaattgggat	ctagccagaa
47461	tttgtcccag	ggctaagtgg	ctaaggcagt	gaactagtct	gtcagcttct	ctgttttaaa
47521	tggaagtcatg	taattccatt	acataccctt	tgctttcaac	tatatggcca	aacatgcctg
47581	ggcagcaggc	tgggggacag	aagacatgtg	ggcttgtgct	tgtaYctgg	aattgaaaag
47641	agctcttggc	atggttcttg	ttcaagctcc	atggttgact	agcacatgac	caccaagaga
47701	gctactgtac	ttctccgaga	ctcattttct	cttctgtaaa	aaatgactga	gtttctttca
47761	cagtgcataa	atcttttctt	ccaggaacag	agactgccca	gggcctgtcc	aagacaacta
47821	ggaaaccaga	gcagcagtta	agccattcct	gggcagggaat	aaatgagcag	atgtcagctc
47881	tcagaatctc	agttttctca	ctcataaaat	ggaaaagaat	actattttct	ttatgtcatt
47941	caccatattg	tgagctgaag	gcactttaac	gatgctgagg	ctgtaaacta	atgcataatg
48001	tgatgtttat	cattcttcac	ctatgtttta	gttactgagc	atctcctgac	accagagagg
48061	acgcaacaaa	aatcacaact	cctagaactg	cacagtctaa	atggcaacca	ctagcccctg
48121	tgacaactga	tcacttgaaa	agtgagttagt	ccaaattaaag	atgtgctgta	ggtgcaaaat
48181	atatgctgga	tttggaaagt	gttctatgaa	aaaagtatct	cattagtcatt	ttttatattg
48241	attacatatt	aaaaagattc	cattttggct	atgctggact	cttcagatat	atcctgggag
48301	attgtattaa	aagaaaatc	atgatggaaa	ttagtgtagg	tagttactag	aaaatgtaaa
48361	attacatgat	tcacatccta	tgttcactgt	tctagacata	taaataatac	tttgagttat
48421	ttttaagcat	gagaagacag	aaatcgacga	ataaagtga	aaagctaaag	attaaagata
48481	cttacaaatc	attgttgtat	tttttaagag	gtcattttga	gcatattcct	taccattgaa
48541	gcccctgcct	gtctggctcc	tctgccacc	ttgttttgct	aatgggtcat	gctcctatat
48601	gtgcatagac	gaggagcggc	tctcagggac	tgaccgtaag	catagtccagc	acagcacgta
48661	tacccatgct	ctgaccccag	gttggcgtct	tgagggtgga	agcattttcc	cagtccttct
48721	ttgactcatc	acagaggtat	aaaacacgct	ctccaacacg	tgattttttt	ttttttttt
48781	ttttgagaca	gagtcctgct	ctgtcgccca	ggctggagta	cagtggcacg	atcttggtct
48841	actgcaacct	ccaactccca	ggttcaatca	attctcctac	ctcagcctcc	caagtagctg
48901	aaattacagg	cacacaccac	cacgcctggc	taatttttgt	atttttagtag	atacggggtt
48961	tcacaatgtt	ggccaggctg	gtctcaaaact	cccgaacctca	agtgatccac	ccatcttggc
49021	ctcccaaagt	gctgggatta	caggctgagc	cactgtgcca	gttaaaagaa	aaaaaatatt
49081	caaacatatt	tggtttctta	ttgaggtgaa	gttcacataa	catgaaacta	acagtttttaa
49141	agtgaacaat	tcagtagcat	ttaattaaca	atgttgctca	accaacacct	ctgtctagtt
49201	ccaaacattt	ttgtcaccct	aaaaggaaac	cacgtgctca	gaaggcagtc	actccacatc

49261	accctctctc	ctagtcctat	gctttgcaca	caactgactc	accctgacag	ttataaccta
49321	aaagttaaat	aaaattatta	gtaaactgca	aaccctggc	aagtgaatta	cttaatatct
49381	cttaaaatcc	tgacatcaaa	tgattccct	gtRaggccat	gagatgataa	gaacatccac
49441	ctccctctct	tgctccagtg	gaatgaaacc	ccctcctcaa	ggagggttac	actcaggcat
49501	aggcccgatc	actcatcata	tattgacaga	tgtgcaggcc	tgagaccccg	ccttctggag
49561	ctttctaatg	ctccttggtt	ctcaatccaa	tccttgctag	gaattaggct	cttctgctga
49621	atgtaaaatt	tttccatctg	cctccataaa	gtatgtgcca	cctgtcagct	gactgggtcaa
49681	ccccaaagctc	aaagtatatg	gacatttttg	gggtgacaat	gggtgggcata	catcttttagc
49741	aaagattaag	aaggtaccag	aggaggacca	gcacaaccca	gcaatgcagc	agaatYcatc
49801	aggaaagcct	tggtctctgg	tgagagctcct	tgcttttagct	ctgcctttga	gcgtctgggt
49861	ggctggggag	gcactggggg	ggaagggaag	gaagcaggaa	ttcgtgtgct	gcaaatttca
49921	ctacacttga	ccacaacttg	gagtggaag	ggcactgtgg	tRcgtctatc	tgcccttccct
49981	ttacaagcct	taaaatgctg	tggaatccat	ctggatggaa	gacgccatcc	aattataaat
50041	ccttaggata	ctttttattg	ccctttaaga	gttcYgatgc	ttgctcttcc	atRcccatat
50101	gtttctgctt	aagaggcttc	attcaaatca	aattggaaga	gaagtccaga	ttacccaact
50161	acaattacct	gagcaaaaa	ccctcaatt	atcttggtaa	tccagataaa	ttcggagtac
50221	aaaggtgagt	tgcttgata	gcactggcta	agaactgctt	tgtctcccat	ttcaccttgc
50281	attattataa	tctacttgat	tacacttcat	ttataacaat	agggagggt	tgttttatgc
50341	cacacattgc	cacttgctca	agtgtaaaat	aaaccttaaa	tgactgatgt	actgctaaat
50401	ttatctcaaa	aaacaatcat	aacaggaatc	ttaaggcaat	aaagtcaaac	aatagaaatt
50461	ttacagaatt	taaaaattat	tttgtgagtg	atttatgcta	aaatacctca	aatttatttc
50521	aaaaacagtc	ttacttaaaa	ctgacttaaa	atatatgtag	gataaataat	Ytgtagtgaa
50581	tttaacgaaa	cttcagcaat	tcaaatgaat	gcacRttatt	cttcttaaaa	acagatatcc
50641	tgaaaatatc	catcaggata	aatagcta	gcattgtgtg	cctaatactc	aggtgatggg
50701	ttgatagggt	cagggaacca	ccatggcaca	tgtttacctg	tgtaacaaac	ctgcacatcc
50761	ggcacatata	tcctggaacR	taaaattaaa	ttaaattttt	aaaaagcaga	tatttggaaa
50821	ttaccaatag	cttcaaagta	aaaatactct	Ncactgtttt	tgtttgtttg	tttattttcg
50881	tttttgtttt	tcctttttga	gatggagtct	cgctctgttg	cccaggctgg	agtgcagtgg
50941	cgtgttctca	gcttaccaca	acctctgcct	cctgggttca	agtgattctc	ctgacttggc
51001	cgtgccacca	tgcccagcta	atttttttac	atttttggta	gagatgatgg	ggtttctactg
51061	tggtggccag	gctggtctcg	aactcctgac	ctcgtgatcc	accgcctca	gcctcccaaa
51121	gtgctgggat	tacaggcgtg	agtcaccgcg	cctggcctat	gcactgttaa	ttttatcacg
51181	acagttataa	atccatcaag	atgtcctgtg	atttttcaac	acaaattact	caaattatga
51241	acttctctgt	gtgtattttc	cactggcttg	tttatgaatc	ctgctgatat	gtctgctaca
51301	atttctccaa	aggctaaaag	atgtttacag	gacaaatgac	atcatgacca	tgtcaggctt
51361	aaatgtgttt	atatcagttg	ggttttagagg	gtcttcaaaa	catggcttta	aaatctcagt
51421	tattatggag	cttaatgtgt	tcctgatgtt	tctggagtMa	tgattttcta	aaaatagaga
51481	aatagattgg	ttaaaagaaa	atgaattggt	ggtctaattc	gtcttatgca	tggttaacga
51541	aatgcatttt	cactctcttc	cgtgataata	cgacacttaa	aactgtgagg	accaatgggg
51601	tgataagaga	caaaataaac	agtaaataga	tactccatgt	gagttactac	tctgtaatag
51661	atagaaaaga	tattttatag	aaaagggaaa	ataccacacg	acataaagc	aatgtcctgt
51721	ggttgctatt	aaatttactg	ttagctgttt	aaaactgcct	taaaaggttt	ttaaatgtgt
51781	cttttctacc	atagattctt	tagacgacag	ccatcaagca	gcaaatacct	cgcaggaggc
51841	cattgtgttt	ctatttgaac	ttctatggat	gtcaagctct	gttcttcaaa	atthagccct
51901	acagatgggt	ggcccaccgg	agagaattaa	tctcactatg	cagccaacac	taccctttct
51961	ggatcttttc	ccagtagtgt	ggattctgct	ctagtccaga	attatgcaga	ataagtatac
52021	ttcatcttag	acaccataat	tcttcaata	tctaaaatgt	atattatgct	ctccaaatcc
52081	tggtcatctc	cagggttaagt	atctctcaaa	aacagtccat	cagataatat	gatttttaaat
52141	ctttatcccc	ttggtcatat	accacaaatc	tcattctatt	taccaatatac	cctactgatg
52201	cataaactac	atatgtatat	gtgtgttttt	ttttaaaaaa	aaagcaattc	cactgtaatt
52261	ttagcagagg	ttatttccac	atagagctat	gtatatgtgt	atatacacat	atctttatat
52321	tatatagttt	tctgtgcttt	ccaggttatc	tacaatgggc	atatatgact	tttatagcca
52381	ggtaattcaa	tattcattca	ttcagtgaac	atcattgagc	atccactatg	ggctgggaag
52441	tgaggatgga	gccacaaaca	agcaaaaata	aaaccctgc	tcttggtggag	ttcccttctc
52501	attggaagag	acgtacagta	aagaaaaaaa	taaaagagta	aagttaatat	cctgtgagaa
52561	ggggatataa	aagtgggaag	gaagacaggg	tgatgtgtga	gaaacacttt	tttttttttt
52621	ttttcagatg	aagtctcact	ctgttgccca	ggctggaggg	cagtggctcg	atctcagctc
52681	actgcaacct	ccacctctct	gtttcaagca	gttctcctgc	ctcagcctcc	caagtagctg
52741	ggacttcagg	catgcaccac	catgcctagc	taatttttat	attttttagta	gagacaggat
52801	ttcaccatgt	tggccaggct	ggtcgtctca	aacttctgac	ctcaagtgat	ctgcctgcct
52861	tgacctcgga	aagtgtctgg	actacagcaa	tgagctgcca	agcccgccct	ggggaacact
52921	attgaataag	gtggtcaggg	aaggcttttac	tgtgaagggt	atgaagaccc	aaagaaagtg
52981	aggaactaag	tcataaatga	tggttagaac	attccggatt	gcgggagcag	caagtgcaaa
53041	ggccctgggg	tgaggaatac	caagaatggt	tgtgtgggca	gagcagaaag	aaccaagggt

53101	aaagcattag	gagatgaggc	tgaagaagta	accgggatca	ggtggttagg	gccctcagge
53161	cctggaggaa	ggctggcagc	gctggagcca	caYggactga	caatgtctag	cagtcacccc
53221	ggctccctgt	gaagaacaga	ccctcgcttg	cagagaagcg	aggacaatta	ggaaagcaga
53281	gtcatgatcc	acgcacagga	tcacgggtgac	ttggaccaga	aagagagcag	tggagctggg
53341	gagtgaactt	tttcagatatt	tctcaggtgg	taccagtggtg	atctgcagag	aatacatagt
53401	taaaggaaat	gtctggaact	aaacatgact	cttcaaactg	gatcctactg	tggcagaaga
53461	gaggcgactt	tacatttcct	ttcccaggat	atcttattct	aatgaacaca	gaaaccgtaa
53521	aatcaagcca	accaacatga	tgaaaatgat	tatataaggt	acacagccca	ggctggattt
53581	gaactgatga	gggaaaaaaa	atctctagt	aaaattcctc	tctctaacca	cagccttctc
53641	tccttccatc	tctgcctcac	ccttgctctt	tctgaagcct	tcctggatgc	tcatgaagct
53701	cYctccttct	ttgactcatc	ctatttttccc	aggacMttta	tcaactcagg	gctcttctcc
53761	cctccccaca	gtcagctcac	tgtgctctga	ctagYtccct	gaaattgtct	atcctcttac
53821	taaccaggaa	gcttctaccc	gctcttggtc	ttctttctcc	accccagtg	caaagaacta
53881	agcactgatg	gagaaaaagc	atgcggtcaa	tactgtctgt	ctctaccaca	aatctcataa
53941	tttttaaaat	ccatttgtaa	tttacattca	ctccttaact	tcctttcatt	cactcctaata
54001	tcacttcccta	tataattctc	tacacagctg	ctgaaaacct	tcctttctgc	tgccctcttac
54061	cccaaaacta	ctctcacatt	actcacaaaag	gttgaccttg	ttttctatWt	acttttttag
54121	agtaaYgaag	ccataaaatg	taagttcttt	caacctggac	ctccactct	tgggatttct
54181	caatagcttc	acctgtattc	tcctcccttca	ctcttgactc	agaggtaaaM	gctcccttgc
54241	tcctcaaatcc	tttagcaacg	taggggtgat	aattgtctct	ttccaaattt	aatccctgca
54301	tctgcgcccc	tgagacagac	actgccatga	tttcccagct	tgctgtcttc	tcctctctgc
54361	atctcctccc	tttctctgtg	acggatatatt	tcctgtgca	agagaaattt	cttctgcct
54421	aaaataaata	atccaaaaca	tctcttgact	ccacttaagc	taacattcaa	tatctcactt
54481	tccttctgat	taacaaaagt	taattgagtc	agtacggtgc	tacaattctc	agacacaacg
54541	ataaacatga	tagagaccat	caacatgcct	tggcattttg	cctgctttct	cttcaattac
54601	ttgcaacttg	gcttctactg	agctccactg	aaaaagcttt	catgacatga	ccacatccaa
54661	tcctcaagaa	actgtcccc	tgcccagtg	tcttcaaaac	cctccaactc	cacagtgtct
54721	cattaactcc	aactcatttc	ctactaaaca	tctcttctca	gctcttcaaa	gtcaaagatt
54781	gcgaaccaaa	tctacagact	tctcctctct	tctcctttgg	gtccaactcc	atccttctcc
54841	atgaaacttc	ttattcctag	aaatgggtgtc	tgcacctttt	aggcttccag	gctYgagcct
54901	agagctcatc	ttactctttt	ccatacccag	actcccaaga	cccagtcagc	tgttttagtg
54961	tctcctctgt	atcgagcact	tactgtcatt	gccctggctc	atctcccaga	gcccacatct
55021	tgtccactg	tctgcaaagc	ctttcctata	tgctggatat	aagtgcacaa	tcctccctaaa
55081	atcccattga	agtttttact	ttttctctatc	ttcttttttt	acttcttttt	atcttctctt
55141	actcacacag	aagtgtctcca	attatgagga	ccttattttta	ttccctttac	tgtcaaacac
55201	aaagccaaca	tgccatcacga	aaagatagta	ctgtgcttcc	ttcagtgcca	ctctagcacg
55261	cccactccag	cacaccttc	ttgtgcgggt	gtatgctcat	gtgatactat	cacaggcaac
55321	acttactatg	caatacgttt	tcacacatca	caatatttcc	acaatattgt	gatatgagct
55381	cattcttcag	gccttgacac	aagcgcccac	aaacctaaaga	gaatgcattt	tcatttttat
55441	gatacacagt	tttctactgc	tccactctgg	ctgccaaggc	attctctctc	caaatacRa
55501	tgtctatata	aaaagaagaa	caatgagga	gagctggagg	tagcgctaag	aaacaacttg
55561	tgtccagctc	caggctctcat	agagaacact	cagaacacgc	ttgttagagg	tatctgtctc
55621	tggcctcagc	ttctcagaag	caattacaca	tcttctgggt	aaacgtggag	tgtgcagcac
55681	agaataccaa	agctgagctt	cccattgtac	ctggttggtg	atctcctcgt	tgacctccct
55741	cagctgaata	tttcttaaga	acagtgttca	attaatccaa	ctatccctgg	tggctggcag
55801	gatgaatcag	ttatgactta	tttcatgttt	gtgggtgggt	gtatgaagt	gtcctcaaa
55861	gagatttttg	ctctctgcag	acagtctgcc	tgtctgtagg	tcacttctct	ccatttctct
55921	tgagtgggaa	tatgagccag	ggtgttgga	gaaggaagag	tcctaaagat	gtctcaggga
55981	atgtcccata	gccattggcc	tggctctgcc	tagataactt	ctgagccagg	tcaccttaaa
56041	gtgctacatg	aaaagtcatg	gatttttttt	ctaacttggt	gatctcaaat	gtcgtgattt
56101	ggcatgaaata	tctagcttcc	catgttgag	ctatggattt	attcttgcca	aaatattggg
56161	gattgggctg	gagaggaag	aaggtagcaa	gatgcattct	gtgttgctca	acaagatctg
56221	catggcctca	accaatgcgt	gtttgaccat	atggatattt	acaaacccca	ccatcaactc
56281	acaaaaatga	agagtgtctt	agacttatct	tcacaagaca	agagggaaacg	ctgcctctaa
56341	ggagtctctc	tgggacacct	gggagagatt	tatttaattg	taaagcaagg	aaaggaaaag
56401	atgaacatg	acctgcaaat	cattaaagag	ggcgactoga	tcataggcaa	gggcttttat
56461	tagtgctaaa	attccacttc	aaagtgtgct	gattcatgga	tgctttttgt	tttgcgtgtt
56521	aYgttgtaga	aactgatttg	attttcatag	acaaaataac	taggcattta	taaccacata
56581	taggtatgaa	agaaagataa	aatgatgcc	ttagccaggt	cagaaatact	cttaaccatc
56641	aaaaatagtt	cctgaatcta	aYaggccaac	aatattccaa	caccagaacc	tttgctacca
56701	tagccacatg	ttttaagact	gcaccactgt	taagttagag	caatttttaa	tacttaggg
56761	ctaaccacaa	ggcataaaat	gcataaaaaa	ttacataatc	ctgtactttt	taataattaa
56821	atgaaaatac	atagactcta	agcaggaaaa	tatctgagat	ccttaaaacg	gtataatcat
56881	aattgttcta	cattggtaat	atgtgtgatg	atatttttct	ccaaaaaaaa	ttctctacca

56941	tttcattctc	tttRatttct	tctagcccg	aaggctctag	aaagaggaat	acctgaaaa
57001	atgcaaatg	tacccctgtg	ctgttcaatg	acagcaaaaa	ttgcagctcc	ccaaccccc
57061	ctcttatgtc	ccctgagcaa	tctgtgatca	cttgagggtg	aggacagtat	attgccccaa
57121	taaactggct	ttccaccaag	taccaccaag	aaatcaaaac	agcttttatt	ttatttgtgc
57181	atcatatata	agttaatagt	tcataattag	ttcaaaattt	aaaattatgt	tctaaaaata
57241	aagctttcta	agtcaatgta	ggcaacRatt	ccaagatcct	tatagttttc	ctctcttaaa
57301	cagagaccat	tatcatatga	tttttttttt	tttttgagac	ggaagctcac	tctgttgccc
57361	aggctggagt	gcagtgggtg	gatctctgct	caactgcaag	tctgcctcct	gggttcacgc
57421	cattctcctg	cctcagcctc	ccaagtagct	gggactagag	gcaccgcgca	ccacgccccg
57481	ctaatttttt	tttttttttt	tttttttgca	tttttagtaa	agaccacgtt	tgaccgtggt
57541	agccaggatg	gtctcgatct	cctgacctca	tgatctgcca	gccttaacct	cccaagtgct
57601	tgagattaca	ggcgtgagcc	accgtgcctg	gccagtccata	tgattttctaa	agagaaactg
57661	aacagttatt	tgtagaaact	ctgcaaagca	tgactttggg	acgggcacag	tggttcacgc
57721	ctgtaatccc	agcacttttg	gaggctgagg	cgggcagatc	acgaggtcgg	gagatcgaga
57781	ccatcctggc	taacatggtg	aaaccccggc	tctactaaaa	aaatacaaaa	aaatttagcca
57841	ggcatgggtg	tgggcgctg	tagtcccagc	tacttgggag	gctgaggcag	gagaattggc
57901	tgaacccagg	aggcgagct	tgcaagtgcg	caagatcgca	ccaccgcact	ccagcctggg
57961	caacagagca	agactctgtc	tcaaaaaaca	aaacaaaaaca	acaaaaacaaa	acaaaaacaaa
58021	aaagccatgc	ccaagcccac	ccctgaccag	tgaaatctcc	gggattgggg	cacaggtata
58081	atgttttttt	aaacatcttc	ctagatgttt	cctgcatcca	gggtgataaa	tcatcatact
58141	aaagccatgc	tctgaaggct	cctgagtagt	agagaagttc	agaataagga	ccccaactac
58201	ttgcacatcc	tgacactaac	tagctgtagc	tgtagaccct	ggacagggag	agatctaaac
58261	agctagtatt	cttacttcac	cRaaatctac	aactaaagat	gtctatccac	gtgttatagc
58321	tttggtcttc	ctaactctaga	tggtcaaatct	acttttctgc	aatttggcaa	agtaaaacac
58381	acatatagag	taatgcacaa	tgctcctgatt	gttggggttt	attatttgtt	ttgcatttac
58441	gcttcttagg	tatgtaacat	gccaaagtca	tgaggagctga	gtcttttgaa	acagaggctt
58501	ctttgggtgc	aaatattaat	tttctatgaa	aaaaagagaa	agtacaaaaga	aaagaaaaaa
58561	cagcagtaat	acagaaatga	atatcttccct	tgtctaggac	acagtgtttt	ttttaaccta
58621	gatttctttg	aacacagact	gacacaaaaa	ttaaaacta	ttgctccatt	ggagaggtgc
58681	aaacctgggg	cggtgagagt	gaggggaagc	gagttagatt	gtcactacac	tagcttctg
58741	ggacatggag	acagccagtc	actgaccccc	tgccctccct	tgcccatgca	tagcctctg
58801	ataggttaca	cggaaaaaat	ctacctcaca	gcagcccatg	gggtggagga	aaaataggaa
58861	cttagttgct	tacattcttc	ctgcatcccc	ttcctcattg	gtcagagtca	ccctgMatgt
58921	ctgggttgtc	acctgtctcc	aaggtgacca	ctccaaaagt	caggttctac	gcccgggtgt
58981	gggtgtcttc	atccaagtac	agaaatgccca	agaggatccg	gaaacctttt	gacctggttg
59041	ctaggtcaca	tgaaagaagg	gccagccctc	ctgggacagg	taacaggctg	gccctgagct
59101	gtggaagccc	tcacacccag	aaggagtcaa	tctgtctggg	tgccactga	gaccaagcag
59161	gagctgaggg	tcttgagta	atagatggag	gcccattgaga	tgtgttcaga	cacatagtgt
59221	gtttgtcaac	tgcaatgtgt	gtgtgcccc	aaaattcatg	ggtttaaatc	cttaccctca
59281	aggtgacagt	atgaggagga	ggggcctttg	gaaggtgact	aggtcataag	gatggaggct
59341	tcattgaattg	cattagtacc	tttcaaaaag	ggacccagg	gagattccct	cacccttttg
59401	tcacagtgtg	aagaccagc	aagaagatgg	acatctatga	atcaggaaga	gagccctcac
59461	cagaacccaa	ccctgctaca	ctctgatctc	agacttccag	cctttagaac	tgtgagaaat
59521	aagtttctgt	tgtttatatg	ccaStcaatg	tatagcactt	tgttacagca	gcccagcta
59581	agacagccta	ttatgtatga	gcagggtggc	ttgagataat	tgagaaaaa	gagactataa
59641	ctacaaccag	ccagagatgg	gatcaYtgca	aatgggactc	gggttcacac	aaggtttgtg
59701	acacacacac	aaaaaaacat	aaggcagata	aatcagcaag	ggttagcagtc	aggaattgac
59761	agacgattaa	tggtattgag	aaacagaata	gataatttaa	aagcaaacct	ttgtacacaa
59821	aagaattcag	gaggtgataa	aagtaacatt	ttcatatcag	ttggaaagta	agttattcaa
59881	tagcaacaga	acaatcattt	attgaggaaa	aaWaactatt	tctttcctgg	ttcaatcttg
59941	gggtgttgta	tggttccata	aatttatccc	tttttttctg	ggttttctaa	tttttgtgca
60001	cagaaatggt	ttttaatagt	ctctgagggt	ttttttttat	ttctgtgggt	tcagtggtaa
60061	tgtccccttt	gtcatttctg	attatgctta	tttggtcttt	ctctcttttt	tctttatcag
60121	tctagctagt	tgtctatcaa	tcttatttat	tctttcaaaa	aaccaacttt	tatttttgtt
60181	ggtcttttgc	atggtttttt	tgatcttcaa	tttcattcag	ttcagccccg	atttttggtta
60241	tttcttttct	tctgctagca	gtaagtgttg	tttgctcttg	ttttttctag	ttcctccagg
60301	tgtagacatta	ggctgttaat	ttaagatctt	tccaactttt	tgatatcagt	gtttagtgtc
60361	ataaactttc	ctgttaacat	tgctttaact	gtatcccaga	gattctggta	tggtgtatct
60421	ttgttttcat	ttgtttcaaa	taattttatt	atttctgtct	tagtttcatt	gtttacccaa
60481	aagtcattca	ggagcaggtt	gttcaatttc	catgattttg	agagatatcc	ttagcattga
60541	ttttattttt	accgtgctgt	ggtctgaaag	tacggctagt	atgatttcag	ggtttttttt
60601	catttgtttg	aaattgtttt	atagccaagt	gcgtagtcaa	ttttgaagta	tgtgccacgt
60661	gtaggtgaaa	agaatgcata	ttgtgtcggt	gttggttgca	gtgttctgga	gatgtctgtt
60721	aggtccattt	ggtcaagtgt	caagtttagg	tctcaaatat	gttaagttag	ttttctgcct

60781	tgatgatcta	tctaatactg	tcagtggagt	gttgaaatct	cccactatta	ttgtatggtt
60841	atctaagtct	ctcactaggt	ccctaaaact	tgatttatga	atctgagtcg	tccagcattg
60901	ggtgcataca	tatctagaac	agttaagtct	tcttgttgaa	ttggaccctt	tatcattata
60961	taatgccctt	ctttgtcttt	tttgatcatt	gttggtttaa	agtctgaaat	tgaacgcgtg
61021	aacagaccaa	taacgaatta	caaaaattaaa	tcagtaatat	aaagactacc	aaccagaaaa
61081	agccctagac	cagagagatt	cacaaccaa	ttctaccaga	tgtataacaa	agagctacta
61141	ccaatcctac	tgaacttatt	acaaaagatt	taggaggaaa	gactcctcct	taacctcatt
61201	ttatgaggcc	agtatcattt	tgatacagaa	acctggcaga	gacacataaa	gaaaagaaac
61261	ttcaggccaa	tattcatgat	gaacagggat	ttggaaatcc	tcaacaaaat	accagtaaac
61321	caaatccagt	accacatcaa	aaagctaate	caccatgatt	aactacacat	tattcctggg
61381	atgcaagggt	ggctcaacat	atgcagatca	ataaatgtga	ttcatcccat	aaacagaact
61441	aaacacaaaa	accacatgat	catctcaata	gctgcagaaa	agacttccaa	taaaattcaa
61501	catcctttca	tgtttaaaac	actgacaaac	taggcattga	aagaacatac	ctcaaaatac
61561	taacagccat	ctatgacaaa	cccacagcaa	cactgtactg	aacaggaaaa	gctggaagca
61621	ttccccatga	gaattgaaac	aagctaagga	tgcccactct	caccactctt	agtcaatata
61681	gtactggaag	tcttagccag	agcaatcaag	caagagaaaa	aaaaaaaaaa	gaaacaaaaa
61741	acatccaaat	gggaagagag	gaagtcaaac	tatctctctt	cacagaggac	gtgatttatg
61801	tctggaaaac	ctcattgtat	ctgtctaaaa	actcctggat	ctgaaaaaca	atttcacaac
61861	gttttaggat	acaaaatcaa	tgtacaaaaa	tcagttagcat	ttctaatacat	caacaatgtc
61921	caagctgaga	gccatatcaa	gaatgcaatt	ccattcaaaa	tagcaacaaa	aaaaatacaa
61981	tacttaaggg	tacagataac	tggggagtga	aaatctctat	gacaagtatt	acaaaacaaa
62041	gatcaaaaga	atcagagatg	acacaaagaa	atggaaaaac	attccatgcc	catgtagtag
62101	aagaatcaat	attgttaaaa	tggccatact	tcccaaagca	acttatagat	tcaatgctat
62161	ttctatcata	ttactaatta	cattttcaca	gaattagaaa	aaactatttg	gaaattcata
62221	tggaaccaca	gaagagccta	aatagccaaa	gcaattctag	gaaacccaaa	agctgaagta
62281	tcacactacc	tagcttcaaa	ctatgccata	aggccacagt	aaccaagaca	gcatagaact
62341	agtacaaaaa	cagacacata	aaccaatgga	ctagattaga	gaaccacgaa	gtaatgtcac
62401	acatctatag	tcacttaatc	ttcaacaaag	tagacaaaac	aagcaatgtg	gaaaggattc
62461	cctatttcaat	aaatagtgtc	tggataacta	gctagccata	tgcagaagat	tgaaactaga
62521	ctcttttcat	ttaccatata	caaaaatcaa	ctcaagatgg	attaaagact	taaatgtaaa
62581	tcatacaact	ataaaaacct	agaagaaaac	ctagaacata	ccattctgga	cattggccct
62641	ggcaaagatt	tcatgacaga	ctccaaaagc	aatggcaaca	aaaacagaaa	ttgacaaatg
62701	ggacctaat	aaactaaaga	gcttctgcac	agcaaaaaga	actatcaata	gagtatacag
62761	acaatgtaca	gaataggaga	aaatatattg	aaattatgca	tccaacaaag	gtctaatatg
62821	cagaatctat	aaaaaaaaac	ttaagcaaat	taacaagcaa	aaaacaactt	ctttaaaaaa
62881	aatgggccaa	aaaaacatga	aaagacactt	ctcaaaagaa	gacatgtgtg	accaacaagc
62941	atatgaaaaa	acgtctcaaca	tcactaatca	ctagagaaaa	gcaaaccaaa	accaccatga
63001	gatatcatga	cacaccagtc	agaacggcaa	ttttaaaagt	caaaaaataa	cacatgttgg
63061	tgaggctgca	gagaaaaaag	aacagttata	tgctgcgtgt	gtgaatgtaa	attagttcag
63121	acactgtgga	aagcagtttg	gagatttccc	aaagaactta	aatagaactt	accattcaac
63181	ccaataatcc	catcactggg	tatataaagg	aatataaatt	gttctactgt	aaagcacact
63241	gcattattca	cagcattatt	cacaatagca	aagacataga	accaatctag	aagtccatca
63301	gtggtggact	ggataaagaa	aaaaatgtgg	tacacaaaca	ttacagaata	ctacacagcc
63361	ataaaataga	atgaaatcat	gtcctttgca	gtaacatgaa	tggagctgga	ggccattatc
63421	ctaagggagt	taacacagga	gcggaaaaac	aaataccaca	tgttttctct	tataagtggg
63481	agctaaacac	tgagcacaca	tggacacaaa	gaatggaatg	atagacacca	gggctgtcct
63541	gaggggtggag	ggtagacgaa	gggtgaagat	tgaaaaacta	ccaattgggc	actatgctta
63601	ttacatgggt	tatgaaataa	tctgcacacc	aaactctcat	ggcatgcaat	ttaccacagt
63661	aacaaacgta	cacgtgtacc	cccgaacctt	aaataaaaagt	tagaaaaggaa	aaaaaaaaaa
63721	ctaaatccct	atctcacatt	atacatctaa	aaaaatccaa	tagattagat	ttaattaaaa
63781	attaaacctat	aaaacacaga	aagaaaaatac	agatgaaatag	ttatataaatt	taggatgaga
63841	acagcctctg	tcaggatgaa	ctaaagaatt	cacaaaaaaa	aatgttaaca	aattaaatga
63901	catgttaaaa	atctggccag	gggtagtggg	tcacacctgt	aatcccagca	ctttgggatg
63961	ctgaggcggy	cggattacct	gaggtcagga	gtttgagacc	agcctggcca	acatagttaa
64021	acctcgtctc	tactgaaagt	acaaaaatta	tccaggcatg	gtggcacacg	cctgtaacct
64081	cagctactca	ggaggctgag	gcaggagaaat	tgcttgagcc	cgggagacgg	aggctgcagt
64141	gagctgagat	catgccactg	cactccagcc	tggctgacag	agggagactc	tgtctcaaaa
64201	aaaaaaaaatt	aagaaactat	atattgaaag	actataaaaa	atagttttgg	ttttttctc
64261	agtaaaaaaa	taatatatcc	attgtcaaaa	attttctatt	ttcaatatta	ttatttaata
64321	gaaaaatttc	aacaatttct	ggagaaaaata	tatgaaacac	atttgacaga	gttttaacat
64381	ccttaacata	aaggatgtaa	cttatatgtt	acatccttaa	cataattttc	tttgttcagt
64441	taacaagttt	ttattgagta	tctatatattg	aatatgcact	atgtctgaca	ctggcaacac
64501	atttgaggaa	ataaagaaga	aaaattcaag	tccacatgga	aaaatggact	aaggacatga
64561	aaaagtataa	atcatcagta	cacttctaatt	aaagttaata	atcaaaagaa	tgcaaaataa

64621	tgtgacctat	caaggctctg	ctattttgtac	tgtcaaatta	taaaaataca	ataacaaata
64681	ttggcagata	ttggggaaaac	aagtaattgtc	atatactggt	atgtgcttaa	actggtaaaa
64741	gtcttcagaa	aggaaattgg	aaatatatgt	aaaaattttg	gacagttgca	tattctttgt
64801	tctagaaatt	ctacttatat	gattagaacc	taaggaaagt	atcatagatt	taatcaaaga
64861	tagttacagg	atattcaaca	tagccctgta	tataacagcc	aagaatctga	agccaacaaa
64921	aaatatttag	cttaaaatgt	atagaatata	cataggatga	aattgcatag	gtcactaaac
64981	agtttctagg	gacatggaaa	aaaaaatgcc	atattgcaga	aatgcaatat	ttgggattga
65041	taaaagggca	gtttcaaaaa	attgaaaaaa	tagcagttac	tatgacctct	tttggcaaaa
65101	ataattatat	gtaattattgt	tcatggaaaa	atagactgga	aatataacga	ttaaaatgtt
65161	aattgtatct	gggttggtga	gattacatat	aattattata	tataatcttt	acttttttag
65221	attgttgttc	acagtgcacat	caagatttgt	catgggagaa	aaaatgaaga	agattattat
65281	tttgtagtca	ccaaggaga	gaagaagtca	catctctgga	gataggagga	gatagggcac
65341	atgagccaat	cacatttctg	cagccttttg	tggatacaaa	aggggtgggc	caagcctttc
65401	caatatccct	tttctctctt	ggtacccaga	cagactatat	ttaccagcct	cctttgcagc
65461	caggtgtggt	catgtgactc	attcctgggtc	aatgaaaggt	agatggaagc	ctggcctagg
65521	aaacaggtct	acatgcaatc	ttccattctt	ctcccaccca	ctggctgaaa	agaaaggctt
65581	ctgaggacac	agagtggggg	gaaggctagg	tagaggctgg	atgggaagaa	agcagggttc
65641	tcttgctcct	ctcatcagag	agctactgag	gagaccacac	gcaccagac	ccccactg
65701	gactgttaca	ggagtggaaa	ataaactttg	tttagaccac	tgaatttga	gagttagagc
65761	tattagccta	ccctgattaa	cgctggggta	gccccacaca	tctgcccagc	accctcagca
65821	gagcctaacc	caatcatcca	aggttaccca	atagagaact	caggcttcca	ctagaggtagc
65881	cgtcacatag	agttactggg	tcttccaggt	tctgcccact	ccaccgtgtt	ttgtgggact
65941	ctccactgaa	ctcacaagc	acatgtattc	agctactcca	aagctatgtg	tactcaggaa
66001	gagaaagtcc	taaatgtttc	tgtgaccttg	tgccactgat	atcatgggtt	ctatccttca
66061	gagtgtctaa	aaaagtgtta	aatctccaga	atagctttgg	aaaagtggag	aaatttgaga
66121	ctttctctat	ttttagtgtc	taatatgcac	tcctgaaaag	tttattttaa	gccacaagga
66181	aatattttcca	tattgaattt	caatatgtca	atatactatt	ttagaataaa	atggaaagat
66241	ccaggacata	gagcagctga	gaagctacac	agttggcccc	tgggcttcac	tcagacagtg
66301	cttctctagg	ctgcaagtca	taagtctttg	tgtcaccac	ttttagaaga	catccccgg
66361	acttctggga	ccaacctgag	gtttgttagga	atctctgagg	atcaatgatc	aaacatatat
66421	tagctttgaaa	aaaactaat	atatgtgtat	gtttttttaa	aaggtttaagt	ttttaacaat
66481	ttagacgtct	tggcattata	ctctgagaac	gttaaccaat	gagaataaga	gtgacaattt
66541	agatgtaaaa	cctcttaag	gtaaaagata	aatcctaacc	acatggggaa	aacctgtgt
66601	tctgtggatg	atttccctgg	atttctcaaa	agactcttag	aatgttctgg	gtatgacttt
66661	gggggcatct	gttctatgaa	gatgggcacc	ccagcatcac	ctgcatatta	acccttcttc
66721	ttgtccttaa	agatgcagaa	ttagctgaga	ttaggaagat	gtacaaagtg	ttctccctgc
66781	acttcaacta	atgaacaaaa	cttttttaaa	attacttatg	attttgaaact	attttatcat
66841	tttatRtatc	ctatggcaac	tggcatttta	tatttaattt	tatcatcttt	tgcttttttt
66901	ctttgatcta	aatcaggatg	ctaagttttg	gggggggggc	ctacagtttt	ggcagagacc
66961	tcattctaaa	aaagagggtac	cagactatta	aggcagaatc	tgctgagggt	ggccccaac
67021	ctggtctggt	gaacctcact	caactcatca	ttgatttcaa	gggcacagaa	gaggcaagga
67081	gatgacagtc	ggagaaatgg	acatgtacca	tggagaaagg	ggaccaacca	caccttctgt
67141	ccaaactgag	atttcccagc	ctgaaacagg	tcaacactgg	ggttagcaag	aaacttttaa
67201	ctgaatagga	gattaaagta	ataagaagat	gagtactatt	attgKtaca	attatattgg
67261	taatagtttt	aagtttcatt	ctttaagaca	gggtgaaatc	cagtccttct	tctcaaggct
67321	cttgtcagca	ttctactttt	ccaaggacat	tcaaaaattt	ctcatccaaa	gtcaaacatt
67381	agtggtgtgt	cctcaaacaa	gaaatagttt	tccatcagtc	tccaaatctt	tgaatagtaa
67441	atggcaaaag	tatggaaatg	caagatggag	tggggcttaa	tgttgccaca	acattcctac
67501	aatactggga	tagtcttcct	tgggaattac	tttctttcct	atcacgtatc	aataaggaaa
67561	gccaaatgta	gaaataaaga	aatctccccg	atttacttaa	cacacggagg	tgacagttac
67621	agcgcaaatt	gaggaagag	taaatgattt	cMacaaatga	acaaaaaaca	cttaaaagcc
67681	acattttaaga	gaaatggagt	cataaaggca	atatttgtga	gcagataaaa	aatcagcaca
67741	gaaaaagggt	agaaatacac	aatgtctttc	atcaagggtg	aaggaaatgt	aaaataacat
67801	taaaagtacc	aactacataa	tttgaatata	aagagaacaa	tagcaagaaa	ggagagggga
67861	agtgatggcg	atatggagac	acagtctctt	ttgcatgcac	agggcagaac	cgggtgcctca
67921	ggctggagg	gtctgtgag	taccccaaaa	tgaarGcaaa	agagcMccca	gcattgaagc
67981	cacagacctt	gtccatgaa	tcctagctga	tgggctggca	cagagagatc	aggagcagaa
68041	caaggaaatc	cataggagga	tcctcccaaa	gtactttttt	ttctggtcca	aaaatcctta
68101	aatctgtcag	aagcaggagg	ttggagtggg	agaattgcag	agcctggaac	acacctgcct
68161	gctgatgacc	cattctgagg	gggttctttt	cacttggcag	caatggactt	ggtaaaaaaa
68221	ataatagatc	cttctccatg	gatcccatct	cccttttttt	tttttttttt	tttttttga
68281	tggagcccca	ggcagaatta	aagaccaact	ctgggaccca	aattttttaga	ggacaaagg
68341	tgatataatg	gccagcacga	tttagtcatt	tagatttcag	gtctaactaa	ctgatgacag
68401	actttaacaa	aatcactgaR	agaaatttaa	cgtcaacctt	aaataacagt	agaggcagg

68461	aatataacaa	atagagtcaa	aatgggtttat	ctaaagggtct	aagcccaaga	aacatccaac
68521	aataaaatta	atctaaaaca	aatttttttaa	taatcattttt	tttttcttga	gacagagttt
68581	cgctcttggt	gccaggtctg	gagtgccatg	gcacgggtctt	ggctcactgc	acccctgcc
68641	tcccgggttc	aagtgatttc	ctgcctcagt	ctcccaagta	gctgggatta	caggcatgcg
68701	ccaccacgcc	cagctaattt	ttttgtattt	ttagtagaga	tggggtttca	ccatgttgg
68761	caggctggcc	togaactcct	gacctcaggt	gatccacca	ccttgggtctc	ccaaagtgt
68821	gggattatag	gcgtgagaca	ccacaccgg	cctcatcttt	tctttaacat	tacaatttcc
68881	catcatgggt	gcacttgaca	gtagtgggac	tgtccaacag	caaagtaccc	catctaagca
68941	ctcaggacag	aagccatttt	aggtcatgaa	aattatacag	attttgtgag	tgaaggcat
69001	ttagttctga	gcagaaagta	gtcttgcaaa	agggaattt	gcaacaacaa	caaaaaagga
69061	caaggagggt	gacaatgtca	agcaaatgat	aaccctgctg	gagcttcaga	aggaaccagc
69121	tacatccctt	tactgttctt	acagatgcta	tggcagcccc	agagggtgtcc	cactaggatc
69181	tcccttcaag	aaagaatcta	gcatgcatct	gaacaacctc	cagctgttca	cacctcaggc
69241	cctgcctcat	cccagacagc	cctgagccaa	tcactgacac	agcagggtac	tagagcctgg
69301	tcatctctgc	ccaacccggg	gcccctccaa	aggcaagctc	tgctccagac	ctcccgttg
69361	ggttggtcaa	actatcaaat	ctgtacctgg	tctgaggctc	tccttgccca	gtcttgcttc
69421	atcctcattt	tatccttcac	aggcacagcc	cccttccccc	tccaataaaa	cactaacatt
69481	cctaattccg	tctcaacatc	tgcttccag	ggaacctaac	taacacaata	ctcctaaaac
69541	caataccagc	aaaaagtgac	tttccaagcc	cctacatgaa	gactcatttt	ctggaaaata
69601	aaattttaata	taattccaat	tacccatgac	gctatttagc	catttcaaa	gaaatgagaa
69661	agcttcccaa	attgtgggtcc	tcatggagtgc	aactgatcat	caaagtgtgc	tgtttacatg
69721	agggcacaca	ataattagg	ttgtatgtaa	catgctactt	tcagattttc	aggttaagtct
69781	aaatgctttt	aattgatgca	agaataataa	gtccttcagt	tgatttatgt	catgcattaa
69841	tatcaatcaa	aaggttttct	aacacatact	ttttaatact	ataaaatgtg	gggttttttt
69901	taatgtcttt	tttttcttat	tcagagccac	atttcttcac	atgggtatgt	ttacctaac
69961	taagggttat	gttggttcgtt	tggtttttaa	ttcagggttg	aggaatcaag	taacaggctc
70021	ttctaaatcc	ttttattgga	tatatgttag	ttttccatta	tcataaaaaga	ctgaggaaaca
70081	ggaacagcag	catattgttt	aaatgattca	caggtaaaaa	tatcttctcc	ctcttggtct
70141	gttttctttc	ctgatgtctt	tgctcattgc	caacattgtc	tattttatat	ataattcttg
70201	caggacattg	tgccctgaag	gtgaaRttgg	acaggtaagg	agatgaccca	acacccttg
70261	tcacacctgc	aggctctagg	tcacttttgt	aacacacaat	gaaatgggta	aattctacac
70321	aatccttttt	tttttttttt	caaacaataa	cagaaagtaa	atatttgaac	tgggggcttt
70381	ttcaaacaac	acaagtcaga	aaaaattgac	ttgtttttga	ctgttacttt	aggcatatat
70441	ttttgattaa	tgacaattaa	atgtaaatat	aaagatatatt	taaacttgag	gttttaaaac
70501	aaagaatgca	agaatcacct	ggttaaagtt	agatatattg	tcagtcctcaa	agtcagttgg
70561	ctggataata	ccatatcata	aaattttata	agactaattt	gaaggctcat	atgaattggt
70621	attattgttc	gtttttcaat	attgatctta	gcagtgtatt	tacagcatac	agcagagtta
70681	atgtttaata	caaacaacaa	agactttcca	gaaaatctac	aaagaatcac	taactaatag
70741	ccatgcttca	gttaaatttg	tttacttttt	aaagtacatt	ctcagaagta	acatgtactc
70801	tcaatttaagY	aaaacacaa	cagaattttg	gaatttgatc	tttaaaaata	taaaaaaact
70861	ttctctgcta	tttacagcct	tggttaaaat	cccaacatct	ggataccaga	atctgagcag
70921	caaaatccac	ggagacacat	gtgcaatttt	atgcctgcaa	gctcactatt	atttagaacg
70981	aaaccaacgt	agtgcattca	cccttggttc	ttgctacaca	gaatgtatcc	agagtacaaa
71041	ctaattgtaga	ttattataaa	actcagtggt	tagatattat	atatatatac	acatgcatac
71101	atacacgcac	atatataact	aaaacaaaat	cttcaaaaat	aatacttttt	ctttctgtat
71161	gtgatgtgcc	ctgatagttt	ccactttact	ccatcccatt	tccttttggt	aatactggcc
71221	ttgactcact	tagacaaatg	ggtcatgacc	tgtagtttgg	aaaacagggc	tttccagcct
71281	ggcttttaag	aggctgggtc	tcaggccagc	agaaccagca	gcacctgggg	gtagtttgtg
71341	gcgcacaatc	tcaaggcctg	ccctagacct	actgaactag	gatctgctct	tagcaagggtc
71401	cccagggtgac	ttatggatgt	attaaagttt	gagaagccgt	cctcctgcat	gcaaaagtac
71461	agaatgagta	tctgtgatct	gaactacatt	taaatcgata	atgcacctct	taaactgtct
71521	ttgttcaaag	aacttctatc	ctactgagct	caaggagcaa	cacagtittct	agcaagatag
71581	gtgcaaagaa	aaattcactt	tcaatttcat	caactgcttc	aagaggttgc	agggactggc
71641	tgacctggta	tcccagctcc	tgacaggaccg	agagtgtggg	aggtgggttg	aggggaggtta
71701	ctgaattgac	tgcaatagat	ccatgagcat	ataaaaagtg	gcgtgtttgt	tcttattcct
71761	gggagaaaat	atatagtatt	cacctaaatt	taaaagagtt	ttatggttta	aaaagttaa
71821	aagagctact	aagccatgaa	aagacgagga	ggaagcttaa	atgcacatta	ctataccaga
71881	taagaaaatc	tgaaaaggct	acatagtgtta	tgattccaac	catatgacag	tctggaaatg
71941	gcaaaactat	ggagatgtaa	aaagatcagt	ggttgtcagg	gcctaggagg	tcgggagggga
72001	tgaacagaca	cagctcagtg	aatttgaggg	gcagtgaac	taactgtgat	gatattataa
72061	tggtggatcc	gtgtcatcac	ctgtttgtcc	aaacccaacag	aacatataac	accaagggtta
72121	tacgggtgtt	atattaaagt	ataaccctaa	tgagaactct	ggactttggg	tgataatgat
72181	gtgttgatgt	agggttcagcc	ctgtttacac	gtcagccacc	catgtgggaa	aatgttgata
72241	gtgggtRgaag	ctatgtatgt	ataggggcag	gaatttatatg	agaacacttg	tggtatatgc

72301	tcaagtctgt	tgtgaaccta	aaactgctct	aaaaaataaa	gtctgtaggt	atctatgaaa
72361	gcgagacac	atgggtccgtg	agagcacgctc	ctctctgcta	taaattgggtt	gtatgggttac
72421	aatcatacct	agaaaggaga	gaaaataaag	agtaagtga	taccgacatg	ttcaattctg
72481	ccccagaata	caacaaagtc	cctgaatcac	tcttgtttca	ttgttgccat	tgcttattca
72541	attcagtc	ctcgtatatg	tgaagcatgt	gctacatata	taacgtataa	tgcaggatgc
72601	aaaaatgatt	atgtgatcac	ttcaagta	ttataatcta	gcaaacatat	aaaacaggca
72661	cacagataat	ccttacatag	taaatgcctc	ctttttcccc	aaaatggcaa	ggagacaaat
72721	agcaaggatc	cgaggacaga	gaggcaaaat	tcagaaggcc	agaaaggcat	ggagccagaa
72781	atccaattca	ctcgcgagta	aggtatcggtg	agagagtgg	gtacaacgtg	acagggacac
72841	ctgttgctcc	caattgta	cttttctgat	tggtttatgc	ccctttccaa	ttactgtgtc
72901	catgctaatt	atcaactatt	tttattatca	accctgcaga	ggggactaag	agaaaaaaa
72961	aaactgctgc	aaagtgatgc	tgtgttcatt	ttgcacagaa	cccagaaatc	ctgatgaagg
73021	aaactgaatt	ggatgccaca	agccttatct	cctgtgaatg	ctgaatgact	ttgtcgggtg
73081	gttttacgat	aagatgtact	ttcttttata	tctatataac	aagctatata	gaatactggt
73141	aaatacacag	aagtaaaaag	aagaataaaa	tagataaaa	tttctaagtt	aaaagctaag
73201	ggaaaaataa	aaagaattat	caattgatgt	gatcaaaaga	aaatgtttat	ctcatgggtg
73261	gttggaag	gaacatcata	aaaattaata	ggtagaataa	tcatcaaaaa	caatctgcac
73321	aacaaatK	acatagctga	ggggaactat	cattaaaata	taatgtatac	aaattgctga
73381	gaaactcact	acaccttcta	aatcatgtca	aaagagtgtg	ttgcatataa	gctgtctgca
73441	cttcttctct	ctctctYttt	ttttttgttc	atattcactg	caaggtgggt	tttgtccacc
73501	catgcaaatc	atcatatga	aggtcgaatc	gatgtccagg	tggccaaagc	ctgtggctaa
73561	tcctcagtc	tcatctcacc	gagacagcag	tttcttagcc	cttctcttac	ctccactggc
73621	ttctcttctc	agctcctctg	tagttcccc	tcctctttcc	aacctgtatg	cattggaggg
73681	gcccaaagct	caacccta	gctctctg	atcttggtcc	ctaggtgatc	tcacctagtc
73741	accagcctc	agatagcctt	aggtcttaga	atcacaaata	tatctctcac	atgacctcta
73801	gactcatatc	aacctctcaa	catcacaaat	tggatatcca	acaggcatct	ccaactcaac
73861	acagccaaac	cccacttttg	acttctgtct	cattgcaact	gtccccctct	gcagatgtcc
73921	cttttagagt	aggtggcacc	atcatatacc	cagttacaaa	caaaatcttt	aaagtcatcc
73981	ttgactctcc	tacactgcat	atctaatacca	tcagcaaatg	ctgtcttttc	caccttcaaa
74041	gtagatcccc	agtctgatca	tttctcacta	cctccccctt	cgctgctgg	tgccctggcaa
74101	tggcctccta	attggtcttc	cctgtttcaa	ctcttggtcc	ccctctagac	tgtttcccac
74161	acaacagcca	acatggttct	tttagaactt	caaccaatgt	aaattatact	tcaagaaagc
74221	tttaaaaaat	tcaatcagat	cacatctgct	atgaaccttc	caatggttct	cctcccgag
74281	tgtgccaaat	aaaatggaaa	tgcccttatc	cttgcttaac	agaagtctac	aacacagcct
74341	caccagcatg	cctccaactc	acatcctgca	gctttaacct	cactccacRc	cagccccag
74401	tctccttgcc	gttacttgaa	agtgccagga	gagctcctgc	caaagggcct	ttgcacatgc
74461	tgttccctct	ctctgcaata	ctcttccccg	tgtgatcata	cggctcatgc	tcccactccc
74521	ttcgggtctc	agcttgaatg	tcacctcctc	cctatgacct	tccctaacct	ttcgtgtaaa
74581	attctacctc	tccttctctc	agtctcttgc	ttctcctgct	ttgttctgga	gaataccNaM
74641	cagcaccacc	tgatttgtaa	ctattccctt	gagcacattc	tatctcagcc	acgggaatgg
74701	aagcttcatg	caggcgggga	catcatctgt	tcagttcatg	ctgtagccca	gcaattctga
74761	atgattagat	tgaatttgga	tccaaaaata	tcacaaatga	gataacgtaa	cttgttgata
74821	agtgtgtgaa	gagtatgaag	agatgatcaa	gctcacaagt	ttccaagtgc	taacatctat
74881	aaaactccat	tgtaacccat	caaatacagca	aagattaaaa	accacaatag	aatgtgataa
74941	aatagtgtc	tcagatactt	cctggagaag	tatagatggg	cagaattggg	tgatctggtg
75001	tcaatgatca	ttctttcata	ccaatacacc	atgggatttt	gcgatcagaa	ggtaaaagaag
75061	acacagaagg	aaaagaagg	gaagattaaa	ccagcaaaaa	cccataataa	gtaggacttc
75121	ttcctattca	ttagtaagaa	agagttgggt	gtccccatgg	gaaaatgcta	atthagtgct
75181	ctgaaccact	gtaccctcca	attcctgctt	ccccacaaac	agtggcccca	cactgagtcc
75241	tacaacgagg	ataatgatcc	tcaaccccca	attgcacaga	cagggtctaa	tcagagacag
75301	tgtatgccaa	cttgctgatg	tctctaaaac	actcaactac	tgggggctaa	ccctaagMaa
75361	aattattgac	tccgattcct	aagttaattt	cctataagg	ggaagaagca	acaagtgtca
75421	cttctcttct	gggtctctatg	acaccaaaata	gtggagccga	agagacttgc	tgatgttttg
75481	aaacccaagg	ccaataaatg	ctctctaaac	ccttagagg	gtgactagga	tattcacaga
75541	agccacgcaa	aagacagcag	ccttcaattga	agaacaaag	aaacaaaaaa	catggttctc
75601	tttataaata	tttatcgatt	atctaaggct	aggttaattg	ttgataacta	tcaattatct
75661	gagcaagata	cgcagaaaca	cttcagggtgc	ttgaactgca	gcctcctggg	ctccaacca
75721	aacttactgg	agaagaagct	ttagatacca	ggaatctgca	ttttaacca	ctcctcagg
75781	aaactgcccc	agagggggctc	taggctaacc	atgcgcaact	tccaccatt	ctcctgggaa
75841	tgtagttgta	atctcctcca	aacctctctt	acacttctctg	attaagttgg	tgccccctac
75901	gtaatgtcag	cctcactgaa	tcttaaccag	gaatcgccag	actggggggg	gNattgtgca
75961	gcatgtctgg	ggaaacttct	actacttctac	taaaacctgc	tgcaacatg	ctgaacaccg
76021	ttccaaaggc	tacaggaact	gagcaccgca	ttggaatctg	gccctggagg	gtgccagca
76081	tggcgacct	gaggggtRca	ggaacctcag	gagaaggaga	aacagaaggt	acagtccagc

76141	tcgagaccac	gcacataaag	aaatctgtag	gcagggtctc	tattattgtg	ccatattaaa
76201	tatcttgata	ttttaaagca	atattatttt	tctgattata	aaagtactca	ttgtagaaat
76261	tttaattgtg	ccatcatcct	gcagaaaaat	gaaagtacaa	aaaacgaaaa	agcagctata
76321	cacggccaca	aagtcacagg	gctcatcggg	atgtgctaac	caggctggaa	actgatcaaa
76381	gctgatttat	tgagaaaaaa	taggcagcat	atcaaagact	tacatcagct	tcatagtaga
76441	cagactttgg	gggcggggga	actctgtctc	ttttgagtat	attctcaaaa	ggataaactt
76501	tcaccttcct	gtgaagcaca	gttactataa	atcagatttg	caaactgtga	gaaatgctta
76561	caactcaatg	agtcttttgt	cagccagtgg	gggctccatc	gtgaaatcca	cccttcgcca
76621	tagcgatggg	ataaatcatc	accagccccg	agacatcgcc	taatcccttc	ctcgcaaggc
76681	agattcgggc	gctttctgca	tctgcttcgt	gttattgaaa	agagctcggc	tcatctgctg
76741	ggctctcagg	cttcacagtg	acctcgcaca	gaacgctgct	acttctgaaa	ctggttttaga
76801	gtcttccgct	tatgaaacag	acacagagtt	taagttagc	gtcttttatg	gtcagaaata
76861	tttagcggtta	cttccagagc	ggcagccgtg	gaacaatgtg	ggaaagaagt	gcttcagttc
76921	acttcttggg	aagtactgac	cctgggtcaa	tttaattttt	tccgtaagat	ttcccttctt
76981	tacttttctc	cttccactgc	tatatgtaac	agtatctccc	cttccactgc	tatatgtaac
77041	agtatctccc	cttccactgc	tatatgtaac	agtatgctag	tatacaacat	tatagattaa
77101	taatatgcta	tattgtttta	aaatattaat	gttatataat	taacatgggtg	tgcatattat
77161	tatattgaca	atacgtaatg	tgatgtgata	ataagtctca	ctgagtgccca	ggcaccgctt
77221	tgaatgctct	aaactcatta	gctaatttat	tcctcgtgat	tctatggggt	agatgcacta
77281	tttcattccc	attttacagc	tgaggaaact	gaggcagcta	gtaacagggtg	gagccaagct
77341	tcacactgat	atcatctggc	cccagcacac	tcataaaaat	gagcaaaagt	ataaaagcca
77401	aagggaaaag	gaagccaagg	ggtgaaaatc	ctaattacaa	acaaatcctg	ggtaagcttg
77461	gttttttagtc	cttccccaca	tgtgaacagc	aactagaaac	ctgaacagtg	actttgtttc
77521	ccctgtcatc	cattgttggt	gcctgatcct	cacacaaaat	cttacggagc	agtatctagt
77581	gattagctta	gctgaattga	gcctggacac	ccttcatgag	ctggaggaca	gaggacgtgt
77641	cgatggaggg	tccaacttca	ggcatctgta	ctcgctgct	acagtgagt	ggacctatga
77701	acccccagcc	tcatcaagca	cccagggtgc	acctgcctgc	ttcacctgga	agagacagca
77761	aggtgtgatc	ccaggaagat	cctctgccag	gaggcttaca	ggcagaagca	gctgctgctg
77821	aacatagatg	ctgtcRgcag	gacacagggg	tgataggggg	tgagccagtg	gagcaggtga
77881	aggtgcagac	tacaggcagc	ccccctgaga	tcctgccag	acccccagcag	cagctccccct
77941	gtccctctct	agcagcttcg	ctgtgcccg	agtgtttgag	aacaggtgac	ccatgccttg
78001	agtggcagat	gctgtgtctt	tttgatccct	tctctccctc	ccccccacc	cccagtgggg
78061	tacatggaag	gtttgctaaa	tgtttgttga	atgaataaac	gagcccaagt	cagaagcagt
78121	ctccaaatca	ttcctcatcg	caaaagggtg	cgttggaana	actcacaaaag	ccaaggcagg
78181	aggacacact	caacttgga	acggcccgtg	cccagctttg	gaacaggcac	agcctgagct
78241	tcaagggaagt	caacttgacca	gcagtagctc	ccaccttcc	ttgcagctag	aacagatca
78301	gaccaagtga	tctctaagt	gatgatttgc	taaagtacaa	tgagttttac	aactagaaaag
78361	taccttcctc	tcaagactga	cttcaactatg	agccacctaa	tttttctct	ctcttctttt
78421	ggtatttttg	ttgcaaatag	tacacacaMa	acaggatata	taatctattt	gtgcaattta
78481	agaataaaga	agcaaacacc	tttgcaKgt	ctccctagtc	cacaacagtc	agccccagga
78541	cctcagaagc	cccttatgcc	tttcccctaa	aactctcccc	tccttgacaa	ggtaattcca
78601	atgctgagtt	atgtggtttt	ttttccttat	ttttcttttc	aacacatatt	acgtttctct
78661	gaataaatatg	cttttttaatg	ttgcctcttt	ttaatcttta	tttgtggaat	catattgtat
78721	gtaatctgtg	acttttttct	actcacatat	gagaaccatc	cacactgata	catggctaca
78781	gatagacact	tggaagtttat	ttttcactgc	tgtatagtag	tccattgggtg	gagtacaata
78841	ggtttatgta	tccacttttc	tgtagatgga	cacttggggt	gttctgtttt	tgctcataca
78901	tacctcctgg	tgctcaggta	caagggaata	cagacccagg	gtgagactgc	tggatcacag
78961	gctaggcaca	tcttcagctt	tactaggtaa	caccaagcca	atacaaaagt	gcttgtacaa
79021	actcatgctc	ctgccagcat	ggaggaccca	ctgtgccaca	taaaaacatc	actgttatta
79081	acttgtttta	tttgcatatg	attatttttt	cccttatccc	caaccacttc	tcaaagcaac
79141	tagccagcat	gtgccacaca	ggcaagcacc	cactccaatg	tgatctatat	ggttaggtat
79201	gtatgtaccc	tggaanaacc	tgtgatgttg	cttgaggagt	atataacttt	tttatgtatg
79261	taagtagcat	gtgctacaaa	gcgcgttctg	ttttctgtat	ttttcactcc	actcagttt
79321	ttgtaaaaaca	tgctcatggt	gctccatagg	tatcagagac	aggatgatgt	tttatcttta
79381	agatgggtac	gccaccctca	ttttatagat	cacagtgata	gctcacatt	gtttagttact
79441	cagtggggtc	taaatgtact	aactcgttta	atctcccacg	aacctgcaa	aacagatgct
79501	tccgttttcat	ataaaatgag	accaaggcac	aaagaaacta	aggaaacttg	acaaaatcac
79561	caagtttagta	aacatcagag	caggatttga	attgagccag	aatgtgttct	tatccacgac
79621	atgaaactgg	ataagaatgc	aaagctcaga	gaattttcat	gactagcatc	tctgggattc
79681	acagctggcg	accacggagc	agaactaata	tacagatggt	tcccctctca	gtccagctct
79741	acttccatcg	gtcccaaat	gtgtattatt	ctatatgatt	agctggtgaa	caacaaaggc
79801	aattagtttag	tcttcttcag	caaagtgttag	taaagtaaaa	tattatagag	acagaactct
79861	tttttaggaaa	tgcaatcatg	tgaattttaa	aatgctttat	cagaaataga	attaatgaca
79921	aaaatatttta	taggatccta	agaaaatgga	tttgctattt	ttttgtcgg	ctgatgataa

79981	tcatctgggt	gcactctttg	ttgtgacagc	tccacccctt	cagagagttg	Katcttccata
80041	cacacctaac	atgctaagag	tagatgtgca	tgatatggta	gtagatataa	tttctgtaat
80101	aaaaaaagta	cttgggtcatg	cctgtaatcc	cagcattttg	ggaggctgag	gtgggaggat
80161	tgcttgagcc	cagagtttga	gaccagcctg	ggcaagatag	tgagaccttg	tctctacaaa
80221	aaaaataaaa	attagctggg	tgtagcgggtg	cacacctgta	gtcccagcta	ctcaggagggt
80281	tgaggcagga	agatcgcttg	agcctgggag	gcagaggctg	cagtgaagca	taatacatacc
80341	actgcactcc	agcctgggca	acagagcaag	actatgtctt	aaaaaaaaata	aataaaaaata
80401	aataataaat	aaaaaatttt	taaaaaagggt	acctggaacg	gcctgcaaag	aaagggaaat
80461	caaatttatgc	ttcaagacag	tcactattat	cactgacaga	gaagcagcta	aaataaacag
80521	cctgctcatt	ttgtttgtaa	atgaattgcc	acaattcttg	cactatttaa	atcaagaata
80581	ttttatatte	catagttttc	aaccacccct	ccttttttta	ttgtttccag	ctccttaaac
80641	atataaaaagg	ataaaaagcca	tttgttctgg	cagaaaagag	tcatacataaa	ttttgcagag
80701	tgaattttctc	taagatgaaa	ggaatgttcc	tctcctccag	gacagcctac	ggacaggact
80761	gtgatgacac	agagtcttca	cggaatcac	acctgcagtg	ttaagccact	tctgtgaaa
80821	actcaagtcc	ttctcagata	acaaaaacat	tttagtagca	atgaaccaa	aacaaaatct
80881	tccactatat	atgttaatac	atcaaaagga	accattcatg	tgccatttat	aagaacttcc
80941	ttccggccag	gtcggtggc	tcacgcctgt	aatcccagca	ccttgaggag	ccaaggcagg
81001	tggtatcacga	ggtcaggaga	ttgagaccat	cctggctaac	acggtgaaac	cccgctctta
81061	ctaaaaaaa	aaaaaaaaaat	gcaaaaaatt	agccaggcgt	ggtggcagggt	gcctgtagtc
81121	ccagctactc	aggaggctga	ggcaggagaa	tgccgggaac	ccaggaggca	gaggttgacg
81181	caagccaaga	tcgcgccact	gcactccagc	ctgggcaaca	aagcgagact	ccatcaaaaa
81241	aaaaaaaaaaa	aaaaaagaac	ttccttccca	gaaatcatag	cagcgttcc	cagcccaagg
81301	ccagaagaaa	cttctcgac	atgaaatttt	acttcaaaaa	ctcccatatc	aggcatttta
81361	atgaatcagt	ggaagaaaat	tatcctccca	actttgttgt	caagaacccc	aatttcatat
81421	gcagaaaaaa	gaattacaaa	aataattatt	cgaattctaa	ttggtgggat	tttgataggt
81481	ttataaagga	aataaatatc	tttaacacaa	ttctacctcg	cagccaacat	cagctattag
81541	caaccacaga	tcttattgcc	agttcatctc	tccaaacaca	cagaagccat	gtgagttgaa
81601	atattttttt	taaaaactct	cagttatgat	ttctcaaaat	tatatattcat	atattgacat
81661	caaccaatat	gatgatgata	ttatatattg	acatYaacca	atagcatcaa	ccaatatgat
81721	gataattgta	gggaaaagaa	agagagatct	gactgttact	gtgtctatgt	agaaaggaaa
81781	gacataaagag	actccatttt	gaaaaagacc	tgtactttta	acaattgctt	tgcttagatg
81841	ttgttaattt	gtagctttgc	cccaaccact	ttgccttagc	cactttgacc	caacctggag
81901	ctcacaaaaa	catgtgttgt	ataaaatcaa	tgtgtaaggg	atctagggct	gtgcaggacg
81961	tgcccttgta	acatgtttac	aagcagtata	ccttgtaaaa	gtcatcgcca	ttctctagcc
82021	tcaataaacc	aggggcacaa	tgactgcgg	aaagccgcag	ggacctctgc	ccttgaaagc
82081	gggttattgt	ccaaggtttc	tccatgtga	tagtctgaaa	tatggcctcg	tggttagaga
82141	aagacctgac	tgccccccag	cccgacaccc	gtaaagggtc	tgtgctgagg	tggattagta
82201	aaagaggaaa	gcctcttgca	gttgagatag	aggaaggcca	ctgtctcctg	tctgcccctg
82261	ggaactgaat	gtctcagtat	aaaacccaat	tgtacatttg	ttcaattctg	agatgagaga
82321	aaaaaccgcc	tatggtggaa	ggtgagacat	gtttacagca	atgctgcttt	gttattcttt
82381	actccgctga	gatgtttggg	tggagagaaa	catacatctg	gcctatgtgc	acatccaggc
82441	atagtacctt	cccttgaact	taattatgac	atagattctt	ttgctcacat	gttttttgct
82501	gaccttctcc	ttattatcac	cctgctctcc	tactacattc	ctttttgcta	aaataatgaa
82561	aataataatc	aattaaaact	gaggggaactc	agaggccggt	gccggtgcag	gtccttggtg
82621	tgctgagcgc	cggtccccctg	agccccactg	tgcttctcta	tactttgtct	ctgtgtctta
82681	ttctttttct	cagtctctcg	tcccacttga	ctagaaatac	ccacagttgt	ggaggggaag
82741	gtcacccctt	cagataataa	gataaatcag	agccRaaaac	atagttcatt	cctgRaatac
82801	aatatagaac	gtttctacca	acacacattt	tcaaagatgg	tgttttcagg	tttgtatggg
82861	attgtttcac	ctttatgaat	tagcttttta	aaacctgtaa	tactttcttg	ctgacttcat
82921	catttccctc	tgggtgcctg	gaggggtggt	gtggaatgct	aataaaatga	atttcagttg
82981	ccgtccctca	caatctgaac	tatagaagat	atttattgtc	tcaacttccc	tatgagcttt
83041	tctttgcac	atgcagcttc	atggctcagca	gacactaaaa	attaaagaaa	gtctttctcc
83101	cacttttact	gtgtagcaca	gacacactga	tcaaataatgt	tcccagcact	tccgaagaaa
83161	ccttatcaca	tccacaagag	atttcctact	tgactttttt	ttgataaaaca	caaccatcat
83221	agcccaacaa	aacatcacac	agaaatcaaa	ttatagtgtc	cccataagca	tctgtctctc
83281	tatggccttg	ggcaaaaattt	agaatctctg	acatccctag	agttaccaca	atctctgaaa
83341	gtgaaggagc	accacaaaag	ctcaagggaat	gcattgaaac	tcttattact	tccaattcca
83401	ttttcaaaaa	aagtaaatgt	gtgtgtatgt	gagctgcaca	tggtccagaa	tgctctacca
83461	aatgtattgt	tttaaaatgt	gtgtacgtcc	atgtgttaatt	gtatgtgtgg	gattttgaaa
83521	ggctttttat	attttattca	aaaagggtaca	cttttcaata	tcacaaactt	tgctttcaat
83581	tataagccct	tcccattttc	aattaattgtt	ttaggcaagg	aataatacatt	agtaaaacca
83641	accacaccac	ttagggaaaa	tgaattagta	tggggaaagt	gaagaactgg	ccatgcagtc
83701	aattactttg	gccagcatag	atatagttga	actaaccagc	acaggtgtag	acagcctgtt
83761	tggataaggt	agcttaagcc	ccaatttttt	ctgcagctct	tatttttgct	aaatatatga

83821	aaactacccat	ctcatttcct	ccctgtagct	attaccacaca	tgtatgtgta	gtgtcaaaat
83881	aaatggccag	agtcctctca	gggacctccc	ttggcacaag	tgcacttcc	tcacctctc
83941	acctccagag	cctcaggaag	tgcaggtcag	aaagcagcca	gggcagggtg	gaggtgatgg
84001	agcgctgagt	gtcggcgagc	ccagataaaa	caacgttcca	acatcggaaa	tccaatcatt
84061	cacctacttg	tctcctgctc	gctaaccocca	ggcaagtata	aaggagatgg	aattataaca
84121	ggagttagaa	gaatttggat	gtgatttcat	atattaacte	agcaaaattt	tttctgaaaa
84181	ggtccagaca	gtaaaatattt	tcagctggca	ggccatatgg	tctttgtcga	acctactcaa
84241	ctgtgccatt	atggagtga	ggcagccaca	gatgatata	aaatgaatgc	atggggctgt
84301	gttccaataa	agcttttctt	atggacactg	aaatgtaaat	tccatgcaat	ttacatgtgt
84361	tacgaaatat	gcttcttttg	atttttcaac	catttaaaaa	agtaaaaacc	attttcttag
84421	ctcctgggct	gtgttgactg	acccttaatt	aactgatag	tgggctgcta	tccatgaaaa
84481	cctaggcaca	gaggaaggca	gagctggctt	cccagccaag	acgggactta	ccttgatatg
84541	agctgctggg	tccgggacag	tctgaatcat	gtccttcagt	aagccagccc	atctaccagc
84601	tgttcagaac	Ntgacggcta	gaagacaaaa	ggaacaacat	acataataat	aaacaaYtgc
84661	atttctgtgt	tttatgttct	agcaggacag	gaMacaattc	accaatggca	aaagaatcat
84721	agctttacat	taaacatgtt	tgccataaag	catcctctag	gtacgtgtgg	tccatgaggt
84781	gactagataa	ctcacttccc	ttccattggg	aaaatcactt	acaatgagcc	caagagaacc
84841	ctagtccgac	aggacagaag	aatctagctc	aagctaacc	acatcacatg	cagcatccac
84901	acatccttcc	acgtgggag	catctaata	gggatccatt	aaaggaagac	gtgggactaa
84961	aacagggcac	acttctctat	catactgaga	aaacaggact	ttgaaatccc	acaaaaactN
85021	atgcgatctg	catcatcttt	tcagcttYgt	tattaagtct	ttagtaaaaca	gcacagctac
85081	cattcattga	gtacttcagY	gccagttcct	tgagcgctaa	gggctccaca	cacattttct
85141	catttcattc	caaccacgac	tctgagagct	gtcattccca	ttgtccagct	gtggacgtca
85201	agactcataa	agggttactga	tttgtccagg	atcacctgg	taataagcga	cagaaccaa
85261	cttcaaatca	gattaggcta	tctgactcca	gagcccaaaa	aaaatttgg	ttgacttaat
85321	ttagMggtat	caaccacat	ttgaaggga	cattttaaaa	caccgcttaa	ataaattaaa
85381	tttaaagatg	cctataatcc	aggaaatttc	cagaaataaa	tacttcatcc	aatatatcc
85441	tttattcaaa	tYataaagca	gcatgttctg	accactgtct	gtgccctccc	cttccaccga
85501	gccagtctcc	aacctccctt	tgctgtcaca	ggccccagg	ctgacctcta	cagatgggct
85561	cagcaggggt	ccttgccctc	tgccacctct	tcagggttgg	ttgacgtggc	cccctctgga
85621	gagcaggggt	gagaggagag	gggtatctct	tcccctcacc	ttcccaccac	cagacactgc
85681	cgcccaacct	ccctgggagc	ttggttatgg	ttctgactgc	agctggattc	tagggggacc
85741	tctcccgta	tggggcccct	ccttccaggc	tccagctctc	agtacctgg	cccagccag
85801	tctctggtgc	tctgcagatc	tcagccaagc	catcagttct	tttttggcac	taaaaatgcc
85861	tttttttRtt	tacattttta	gaagccatca	cttgcttttt	atcacctgta	gttttttgtt
85921	tggtttctcc	tttacRatgg	ctctttgcag	taacaaatgc	aaaataaaaa	taaacatata
85981	gattacagaW	cttgagacat	tttcgaaggc	tgtctttcaa	atltgctact	gagcgcatta
86041	ctgtagtgc	aagtgaagc	attcacagga	gtcctcccca	cattcacatc	tgtcctttct
86101	gacctctcac	ggtacaggca	cacctattcc	aacccaatYg	cctctcatga	agagcatcta
86161	cctccagatc	tgcatgctcc	acatctgtct	tcagctctcc	tggaggcctc	ccaagattgg
86221	cggagcttgg	ctcaccaaga	agatgacatg	ctgggaacct	taggcatttt	ttttctataa
86281	ccaaatttct	aggtcttgg	aaatatgtac	ttccccccac	ccccaccccc	aggcctaag
86341	tccctctggg	ggatgagagt	tactgaaagt	acagaggag	tgaagtgcag	tccgtgaaat
86401	gaaccgggaa	gcagatgaag	gtgaccaggc	tcctggtgct	ctcgtgggag	aagaggctgg
86461	gagaggatgg	ggcctggctg	aaggtttctg	ggagcaYagg	aatcctcaac	tccccctatc
86521	ctaagagcag	caggtgccac	aggaaaacac	cactctgctc	ctggggctcat	tcagcccga
86581	aggggagcgt	ctgcaagcca	tcctcacaaa	gggagggcgt	cccactgcac	aacctgctc
86641	tggggagcac	atggctatct	ccgtatgtct	ggacattctc	tctggactag	ttcactccgc
86701	acagttcctc	cctgcctgca	cacacatgct	gacgtccaca	ccagggactg	aagctcgtcc
86761	ccctccctcc	aaacccaggc	tgacctgtgg	ctgcctctga	cctgcagaac	acagcagagg
86821	gaaggtttca	ggagtcccag	gctcaggcct	ttatgggaagt	gggagcttcc	ccttctgccc
86881	tccgtgaaacc	aagtcccat	accctcagca	gcacagcaca	tggcagagga	ccgagtcccc
86941	agtggggcag	ccagcaccaa	cgccaggccc	gggagggagt	tgtcctggac	agtccagaga
87001	actgactctc	agatgaccac	ggcccctgcc	gacatcaggg	aggcagcttc	cagctgagtc
87061	cagtcagccc	acagaatcat	gaccaataat	aaatgggttg	ttgtactaag	ctgctaggat
87121	ttgggttagt	ttattacaca	cacaggtcat	tgaaacaccc	tgtaagtgt	ctagaggcac
87181	ggagaagata	gagggtaaat	agcttgtttc	actgtcacta	tggataagta	caaagRtgg
87241	tccaacaYtc	aactcctcct	gtacccatgc	tttttgcac	ataactttgt	gattcttccc
87301	attaaagagg	tagaattact	tcccaacctt	tatttttggg	cggcctatat	gatctgcttc
87361	agccaacagg	cagtagcaga	aagacaagta	tgacagatcc	tagctgaagc	ctcagagggc
87421	aggacatgtt	cccttttgct	ttctctgaac	tttccaccag	catgggaaca	tggccagatg
87481	catgcccggg	gatgagaggt	acctgaggca	gagtcaccca	ccccaatcac	ccSaccaaga
87541	ctacgtttga	tcatcccata	gacaacaaa	accccagaca	tgggagggag	gcaaccagac
87601	cagccgctag	ccaagtccaa	ccagaatggc	caaccacag	gctcatggct	gaataagtc

87661	ttactgcttg	aagccactga	aatttgggga	atttggtaca	cagcattttt	gtgacaatag
87721	attactgaca	tacactttgc	aatgtacaat	cagattaccg	agatcttcgt	taaaattttt
87781	ctagccagt	ataaaaggag	actcaaacag	ggcatttcaa	aggaagccac	ttcatcccaa
87841	cacagagaca	taatatattcc	actcacctac	gagacatgat	ctcaaaccct	aatcctgccc
87901	tcaaaatggg	tYtggaact	gaaggataac	caagaaacct	ccaccagaag	cctccatggg
87961	attcccagag	cccttcccag	gagcaactga	attctcactt	cacatgatgc	cctgccagcc
88021	tcccctggat	tccaaaaatc	aaaacatcct	tggatatcca	agaccgcagc	aacatcatag
88081	aagttacagc	agccactcgg	tctgtaaMca	accaagacat	aaactaaatg	agcaactggt
88141	tgagtctctt	gtttcttatg	gggaagtttt	gagtactcag	gttgcatcaa	gtataaaatg
88201	acaggggaag	gaatagatag	agattccaga	tgcaacacag	tcctgatcct	tgagggatcg
88261	agatgggaaa	cagatctggt	tcatacagat	atgaagcggg	cagtcctgag	agcaggcagt
88321	tagtggagg	gcccagtgag	gcgtgtggat	atacctagag	gaaagagaga	gctgattcat
88381	gcagagacaa	gaggggtgcca	ccaaggagcc	ttgatgaaca	agaagcatga	ggactggagg
88441	aaagaggggtg	aaaggcatga	agcaggggaa	ctaagtttct	cccagctctg	ccaMattaga
88501	gaccaagggc	gtatgagcct	ctcacaccgc	caccaccaga	gtgatcacac	accaaaggga
88561	ttcactcagc	acgcagcgtg	atggcactgc	caccctccac	accagctcag	cgaaaaccca
88621	acggaccta	agaccaggag	aataaaaagt	caaccaccct	cccctgcctc	ctcaaccac
88681	acggatctct	gactcttttt	ttttttttga	gacggagtct	ctctctgttg	cccaggctgg
88741	agtgcagtgg	tacgatctcg	gctcactgca	acctctgcct	cccgggttca	caccattctc
88801	ctgcctcagc	ctcccaagta	gctgggacta	cagccgcccc	ccaccacgcc	tggttaattt
88861	tttgtatttt	tagtagagac	gggtatttca	cgtgttagcc	aggatggtat	cgatctcctg
88921	acctcttgat	ccaccgcct	cggcctccca	aagcactggg	attacaggcg	tgagccacca
88981	cgcccggtta	attttttgta	tttttagtag	agacggggtt	tcaccgtgtc	ggccagatgg
89041	tctcgatctc	ctgacctcgt	gatccaccca	cgttgccctc	ccaaagtgtc	gggattacag
89101	gcgtgagcca	ccacacctgg	ccctctctga	ctctttttgc	agccatttga	gtgectgttc
89161	ttttttacca	ttcagagggt	attttatgga	ttttttaaag	agtacttcta	atttactttt
89221	gtgttcctaa	tggtcatttc	tcaaatatgt	agtagttggt	aaccaaaca	actaaaacca
89281	tactaaaaat	ttttagtagt	tggtgatagt	ttcatgtcac	tgtaggaacc	tttttatttt
89341	ctattttacc	tgcccttgaa	cccttgccga	acttcactct	atgttgattg	aaaatcattt
89401	ctttaatact	tgtttcttgc	aataaaacta	tttgctaagg	tatatacttc	catatgattt
89461	tgtgtttcat	gacccagcta	tattttatta	ttgggtatct	taaagaaaac	aaaaaggcct
89521	taacaataac	agaattctac	ttgtgtcact	tctttgtttt	tgttttcaaa	atgccaatgt
89581	cttactaaat	agaaacaMac	ttgatttgaa	ttacgttaag	tctaaccaca	taaaaaagat
89641	ggactgggtac	caaaggtaga	agggaaaata	agcattttgt	aattatggag	gcagtattgc
89701	aaaatgggtta	agagaaggcc	taaatatttg	cagggctggt	gcaagaaaat	agatgggaag
89761	ttgagacccc	tcgaccctac	ccctcctctc	ttctcaaacc	caattgtgtt	ctgcgggtta
89821	ggggatgcat	atagtcacgt	ggacatcccg	atccacatgg	ccacactcca	ttcagcaatg
89881	ctccctgctc	atctttcaga	cccagggaaa	cacacacaag	atcttcccaa	gaagatgaac
89941	tccaggaaac	aggctgggtg	agcccaaaag	atgctcagtc	aagtttgga	aggcactctg
90001	ggtgccccag	cacctggggc	atgtcctgga	atggggatgg	gttcagctgg	ggagggcagg
90061	aatggggcct	agaaggcata	ggcaggacgc	tgggctgggg	gcagtgtctg	atatgagcac
90121	agtgcagagc	cagatgaggg	atggaaatca	gatctgccac	ttactccttg	gaaacctata
90181	aatgatcact	ttatcctcca	agagccttag	ttacacattt	gcacaatgag	gataWtatta
90241	ctgtagtata	ataagaactc	catttgccct	tgtctctcgt	ttctggcatg	gagctcctaa
90301	aactctggga	atttcccaag	tgatgggagc	atcttttgtt	ctaagaaggc	agcaggttgg
90361	tgagtcctta	gatagettca	ggatgggggc	taattgcaag	aaagagttag	tcttgattag
90421	atgctcagaa	ccttcagccc	catttcccaa	cctctcagga	gaggagagga	gctagaaatt
90481	gagttgatcc	ccagtggcca	gtgatttaat	caatcctgcc	tatgtaacaa	aatctccata
90541	aaataccctg	acgggggttag	gacagcttct	gggctggtga	gcataattgt	atgaagagt
90601	tggtgcaacc	caatgctaaa	gggacagagg	cgccagggcg	cgggtggtca	cgctgtaat
90661	cccagcacct	tgggaggccg	aggaggggtg	atcacgaggt	caggagatcg	agaccatcct
90721	ggttaacacg	gtgaaacccc	atctctactg	aaaatacaaa	agattagccg	ggcgtgggtg
90781	tggacgcctg	tagtcccagc	tgctcaggag	gctgaggcag	gagaatggcg	tgaacccggg
90841	aggaggagg	tgcagtgaag	cgagatcgtg	ccactgcact	ccagcctggg	caacagagcg
90901	agactccatc	tcacaaaaaa	aaaaaaaaaa	aaaaaagggc	agaggctcct	gtacttggaa
90961	tccttcagaa	ccttgccctg	tgcaactctt	catctggttg	ttcattcata	tgcttaataa
91021	taaactgtaa	tcataagagc	atagcgtttt	cctgagttct	agcaagtgtg	tgacactaaa
91081	agggggggtg	tgggaacccc	caatggtgta	gccatgtcag	acagagtgtg	ggtaacctgg
91141	ggaccaata	cttgtaactg	gcattctgaag	tgaggacggt	gttggggagc	tcagccctta
91201	aacccatgtg	ctaacttccg	gtagtcatgt	tcagaactga	attgaactgt	tggacaccca
91261	gttggtatca	gtgttgaatg	actggtKgtg	actggcggaag	acatgatgtg	tttgggtgtg
91321	ggaaacatga	ataaaaaaatt	tgtaaaaacct	ccaaacgaYg	ttcagcaatt	cacaaagggtg
91381	tcagtgcag	agttggcctg	tatctattct	aactggtcag	tatgcaagtt	gaaccattgt
91441	ttgattatct	cctctgtcta	caccatcagg	ctttgtgag	gattaaacta	aattatgcct

91501	cagaaagaaa	actctctcaa	ggctaggcac	ataatcaaca	ctaaacaagt	ggtaatcatt
91561	ataattaaga	aagtaaaaat	atttcaaatg	gacgattttg	atactgtata	aaatgttgca
91621	actttttttt	ttttgagacg	gagtcttgct	ctgtcgccca	ggctggagcg	cagtggcgca
91681	atcttggctc	actgcaacct	ccacctcccg	ggttccactg	attctcctgc	ctcagcctcc
91741	cgagcagctg	ggattacagg	cacgtgctac	catggctggc	taattaaaaat	gttgcaactt
91801	tttaggaaaa	ataatcctca	ataaataaat	tgctaatttt	cataagagcc	ctttgtcaac
91861	tgactaaaaa	catccttaat	ttagaaatgt	attttttagt	tttaagataa	ctagtcatta
91921	agtgaaaaca	ttgtaggcaa	aatggctcta	agcaatttag	tattcatagg	aatgaatcca
91981	ttccttgaaa	gcattggctg	ctgggtgcac	ttgtcctgga	tgagcagata	ttaggcaaatt
92041	cagtctttca	ccagcaaacc	ctgcccataag	aactgcacag	gcctgattcc	tgccaagcac
92101	accaggcccc	tgactggggc	agaggcctcc	cttatggggc	aggaagaagg	aggcagggag
92161	aaacataggg	cccatgattt	aaggaggctc	tccctctcag	gctcatgcaa	gtgcaggatc
92221	agtacctaaa	ggggagcacc	gcccgaatc	ctgtgcccc	gtcaccttaa	cccagtcacg
92281	ccctggctcc	tgactcctca	agctaagcag	ggaataccag	cgggagacag	aagacagcat
92341	gagaaactga	atgcatctcc	cacatttatc	taaaaaataag	aaaMtggact	tcttcagtag
92401	gaaatggctY	atctttccaa	atgaataaaa	atacttatag	ctatgggttg	aattgtatca
92461	attctctccc	ccagttcatg	cactggagta	ctaaccctcg	gtccctcggg	atgtgatgat
92521	agagaggtag	ggcctttaca	gaggcaatca	aattaaagtg	aggctcattag	tgacctaac
92581	caatgtgact	gatccactta	caaaaaaagg	gacaagattt	aaagacagac	ctgtatagaa
92641	ggacgatggg	gtgaagacac	agggaaaaaga	caaaaagagg	ctggaacaga	ttttcctcaa
92701	agccctcaga	gggaaccact	tgggctgaca	ccttgatctt	ggacatccag	cctccggaac
92761	ggtagagatg	tgaatttctg	ttgttttaag	ctaccagtt	catggcagtt	tgttacaaca
92821	gcactgggaa	aggaaatcac	ctctatggag	ggacttacga	aagtcactta	actaatgcag
92881	attagaggcc	tcacagtcAR	ggcatggcct	caagagccac	atccatttta	attgtaccat
92941	ttctcactga	actctgaccc	tggtaaggcc	atgtctcagg	ctccatggac	tactattcct
93001	tccagtatc	gacccctggc	tcattctggc	aaggagggga	ctcatacatt	ctgcccctg
93061	catgagtgtc	atgtgagcac	cccataaacc	ttgatccctg	cgggtgctct	ccaggggagc
93121	cctagcattc	cagctggcct	catgggcattg	ccttagagag	agggagctga	ccctctcctt
93181	cttcaactgg	attttatttt	tatttttttg	catgttttgt	gaatacaaa	attcactgcc
93241	tgggtgtgcc	ccttgaaaga	aggtttcaat	tgcatcagg	tgaataaacc	acaaaatgac
93301	tgtaagccaa	atgtagtagt	atgatagaaa	aatggaaata	tcgtaaaagc	accagtcgt
93361	agacacctag	cttctggttc	ctctggaaaa	gggagaagag	agtagacagt	gatggagaa
93421	agaggggctg	tgaagactat	agaactgaga	aggcagtaat	tctaaaaatct	ccacactgta
93481	gggccaggaa	caggcagtg	tgggcaatgg	ggcatcttat	ctcagggag	tcaagggtgt
93541	catgtagaac	caacataatt	catccaagta	cagcaaaaga	tccagaaatg	gtctgtggag
93601	gtttggaacc	atgaaaatca	aataggaaag	atacaaaata	ttacatatc	cacatttttt
93661	aaaaggggtc	ctgactaggc	agcataagat	tatttgggtt	ctgttcattt	tacctttagt
93721	tgcccttggg	tctgccatct	tttttctctg	tgagtcattt	gtcttgcttt	tgggtcaacac
93781	ggctttcctc	gggtctccaa	agatcctgga	ataacctgaa	tataagttga	agaatatata
93841	accttatttt	gttttcaaat	attaatacca	gggatcaact	gtacatgtac	cgacaaaaac
93901	tatttcttaa	acattatttt	aaaatatcct	aaaaggggaa	agaacttaaa	tccttcccct
93961	taaaaaaaga	ggggacaaca	ttcctctcct	ctgagggaag	cccagcatta	tgaagcaaaa
94021	atcgggagtg	aaagcaaggc	caagcctaac	tggcatgctg	aagcacacag	atgcttttgc
94081	acagaaaata	ttttgaaaat	gttttctttt	ccttatagcc	cccagctttg	gtaataatgg
94141	cttccctaga	tcattctgtt	agcactttct	gctgctcaga	acagggtagt	gggatttaga
94201	atttagaaaa	catcccgaga	actgtaatta	agcaattgtc	atagtccag	aaaagaaaca
94261	acccttttac	tttgtaaattg	ctaggaggat	ctgtattttt	taaaatctta	gaacactaac
94321	tagctaaaaa	gatagaaaga	aaaaaatctt	tcttacaata	tgccgttagc	acggctccaa
94381	aatccagtc	ctctctctct	tctttttttt	tttttttttt	tttttttttag	atacaggaat
94441	ttcttacaca	actttccaag	taaaatattc	aatccccaaa	taatataatt	agtaRaatcc
94501	ttacaacatg	aRaaattgta	aatggcatct	gtttccgttg	aaaatatttt	taaattagaa
94561	aacatcttaa	gcactgcac	aaatcaatat	ttgtaatggg	cagaattatt	tctaaatggg
94621	tccacaggcc	agttgacatt	cccactctac	ttaggaagcc	cccagccaac	tccacctgaa
94681	gtcttcgaaa	cctatgcata	gaggcaaatt	gcccagatgt	cacagaaaag	catttggtgt
94741	gatacagaaa	ataaaacaca	ttactggggg	aaaaaaatca	gggtactaaa	ccacgtacat
94801	aatatcatcc	caatattgca	caaagaataa	aaaataagtg	gatgaaataa	aagaggggag
94861	gcaggaaagg	aaagaggatg	acctgagaaa	agttatcaaa	atgggttaaca	gtcattggct
94921	ttggctttcc	agtgtgaaat	aggagtctct	atcttctctt	ttttaaaata	tttttaaaat
94981	tattttattt	ttaaaccaac	taataaaaagt	aatatatatt	taccgtgtac	aacattatgt
95041	tttgaaatat	atatatatat	atatatatat	atatatatat	atatatatat	atatatatat
95101	atatacatgt	tggaaatggc	aaagttaagc	aattaacatc	tgcatcattc	cacatactta
95161	tttttttgtg	gtgcgaacac	ttgaaatcta	ctcttagtga	ttttcaagaa	tataatgcat
95221	tggtattaac	tacagtcaat	agatcatagg	tctattcaag	ttcaagtcac	agatctcttg
95281	aaattactcc	tcctgtctaa	ctgaaatttt	gtgtcctttg	accaacatct	cccagttcc

95341	cctacgctca	acatcactaa	tcatacaagga	aatgcaaate	gaaatcatga	tgagtgtcat
95401	ctcacacctg	ttagaatggc	cattatcaaaa	aagacaagtg	ttgggtggga	tgtggaaaaa
95461	gggaactctt	gcccaccgtt	ggtgggaacg	tagatcagta	cagtcattat	ggaaaacagg
95521	atggagggttc	ctcaaaaaca	caaagatcaa	ttacaaaaca	aaatactata	tgtttttaca
95581	atgtactaca	actcaatttg	gtatgtatta	caaatacaaaa	tactacatga	ttgctggcaa
95641	tcccactctg	gagatagatc	caaagaaaat	ggaaaatgaa	atcgggtatgt	caaagaggta
95701	ccagccctcc	caaattcatt	acagcattat	tcacaacaga	caagggtatgg	catcaacctt
95761	agtgtccatc	aacataaacg	cccataaaga	aaagtgtgtc	acatacacaa	tggaaactta
95821	ttcagcattt	agaagagaaa	tcctgtcatt	tgtgacaaca	tgaacaaacc	tggagaactt
95881	tatgtctaat	gaaataagcc	aggcacagaa	agacacatac	tgcattcctct	cacttataca
95941	tagcatctga	aaagtgtgaa	tcctatttcc	ttctttatag	ttttgtgtat	tttccaaatt
96001	ttctaaacca	aaaataattca	tcctttataat	caggggagaa	aactaaggaa	taatcatttt
96061	ataatcctaa	acatgctttt	aaagcattta	tttactgtga	atgtattctt	ttcccatcca
96121	gtcatacatg	ggcccaagga	gctggccttc	cctgctcggc	agcctgatgt	tcaaggggat
96181	gggggagggg	gacggtgcag	tagatcatgc	tttgcaaagc	tgtgcaaate	tttttttttt
96241	ttttttgaga	cggagtctcg	ctctgtcgcc	caggctggag	tgcagtggca	cgatcttggc
96301	tcactgcaac	ctccgctcc	caggttcaca	ccattctcct	gcctcagcct	cccgatagc
96361	tgggattaca	ggcgcccgcc	accacacccg	gctaattttt	gtattttagt	agagacgggg
96421	tttcatcgtg	ttagccagga	cggctcctaa	ctcctgacct	cgtgatccac	ccgcctcggc
96481	caaagtgtctg	ggattacagg	cgtgagccac	ggcgcccgcc	caaagctgtg	caaatactga
96541	gttaataaac	atgttcccca	aatccaaaag	ccaataattga	ctattgcttg	gttaactgtt
96601	ttgcatggat	aaagtgttaa	aggagtattt	ggtttccaat	ttctgcaaag	ggaataaaaag
96661	ttttaaaagc	gtgaactaag	atttccctata	ctatgctatt	ctatagtagg	tttttatgcc
96721	aaaacctcta	ctaggctgaa	aggataattt	gatttatggt	tgagacattg	ttgggcctga
96781	tttaaagccc	agagtcaagc	ccagagagga	ttctcttaaa	tatgtgcaat	gggtattttt
96841	aaattagcat	tatggcggtta	ttaatatttc	atttatcctg	agagcattta	cctgacaacg
96901	taaattatgt	gtaacacagc	tgagaagcca	caaaggatcc	aaaaagatgt	cacgaagtcc
96961	acgaagagaa	gctttgccag	agaaaaggac	actaccagcc	gcgtgagctg	gcctcatcaa
97021	agcaaaaaag	tcagaaaatt	atgtaaacca	gactcaagaa	ctcaaaaata	cagtcgtttg
97081	agaacatttg	taagagtgtca	aactgtctac	cccagctcaa	tcacccctgg	tccagcaacg
97141	gcagaaagcc	aaggtcaagg	agtacatgga	cctgaagaga	cacccccac	caccaccacc
97201	accctccaag	ccctgcccct	ggcagtgagg	acaaaccaca	tgacccctgg	gattcagtag
97261	tcagtcaggt	tctgtgttct	gtatggggag	gtagaagaag	gcaccagatc	atcaggaaat
97321	ctggactgga	aatggccaaa	gctgaacaca	tctgagtacc	tgaagggtaa	gcttaagggtg
97381	tcaaagcagc	catgggtgagg	cagaaagaac	cttgagcttg	gggccccaaag	tttatgctga
97441	cccactagca	gccaacaaaa	cctccctcag	cctcagtttg	ctcatctgaa	aactgtgagt
97501	aagaggctct	gactccagag	aggtggtgtca	agttcaatga	gaagggggtg	ctgatggtcc
97561	agatggacag	atgcaaagca	ctgctttctg	tttttgtctt	tgttttgata	caggattctc
97621	tctgtcatcc	aggctggtgt	gcagtgggtg	caatagctca	ctgcagcctg	gaccctccct
97681	ggctcaagtgt	atcctccacc	ctcagccacc	caactagctc	agactatagg	aaaaagagtc
97741	atgcaccacc	atgcctggct	aatttttaaa	aaaaaaNttt	tggagacatg	agatctcact
97801	attttgtcca	ggccaattct	gaactcctgg	cctcaagcaa	tcctcctctc	tcagcctccc
97861	aaagtgtctg	gattagaggg	gtgagctacc	atgccagca	aaggcattgt	taaattcagg
97921	aatgaatgac	caccttggtc	cccaaaacac	tagacaagag	aaccattctc	tacgcacaga
97981	tctccctggc	caccacgtgt	aagcagcagt	gaatattcca	taggtggaac	tggctactgt
98041	gaagaaagaa	aactccagaa	aagctaccct	cactccaaga	agtagcctgg	acccacgtg
98101	cgtgacagca	gcctggggag	aaggtctccac	tcccaccagc	aatgccaaag	acagagcagc
98161	agcagccgtc	catccccaca	gttatggtga	acacagccac	agagcagcct	gctttttcca
98221	tgtatagcca	cacacaggtg	gcaatggcag	cagacgtaac	tggctgccct	tcctccctta
98281	tggaaagaaa	cataattttg	cttgggtgtt	ctgcctgccc	ccacatgacc	caYggaggat
98341	gaccagcccc	aacaccagtg	taaatcctat	cactgggcac	catctcttct	caccacggat
98401	tgggttaggg	aagggcacat	gatataaatc	tggtcagctc	aacactcggg	gaagcgagct
98461	agaggacttc	taggaaaagt	gttggctgat	gaagaaaggc	catagagtag	tagtctctgc
98521	tgtgtggcac	ttggaactgt	gagagccatc			

[0269] Following are cDNA sequences for *PSMB1* (SEQ ID NO: 5), *TBP* (SEQ ID NO: 6), *PDCD2* (SEQ ID NO: 7 and 8), *ELP3* (SEQ ID NO: 9), *CHDC1* (SEQ ID NO: 10), and *ERG* (SEQ ID NO: 11 and 12).

PSMB1 cDNA sequence (SEQ ID NO: 5)

NM_002793 Homo sapiens proteasome (prosome, macropain) subunit, beta type, 1 (PSMB1), mRNA

```

1 aaggcagcca tctcgccgtg agacagcaag tgtcgcgag ccgtgcgatg ttgtcctcta
61 cagccatgta ttcggtcctt ggcagagact tggggatgga accgcacaga gccgcggggc
121 ctttgcagct gcgattttcg ccctacgttt tcaacggagg tactatactg gcaattgctg
181 gagaagattt tgcaattgtt gcttctgata ctcgattgag tgaagggttt tcaattcata
241 cgcgggatag ccccaaatgt tacaaattaa cagacaaaac agtcattgga tgcagcgggt
301 ttcattggaga ctgtcttacg ctgacaaaga ttattgaagc aagactaaag atgtataagc
361 attccaataa taaggccatg actacggggg caattgctgc aatgctgtct acaatcctgt
421 attcaaggcg cttctttcca tactatgttt acaacatcat cggtggaact gatgaagaag
481 gaaagggggc tgtatacagc ttgatccag tagggtctta ccagagagac tccttcaagg
541 ctggaggctc agcaagtgcc atgctacagc ccctgcttga caaccagggt ggttttaaga
601 acatgcagaa tgtggagcat gttccgctgt ccttgagacag agccatgcgg ctggtgaaag
661 atgtcttcat ttctgcggt gagagagatg tgtacactgg ggacgcactc cggatctgca
721 tagtgaccaa agagggcatc agggaggaaa ctgtttcctt aagggaaggac tgatctgtgt
781 gctcttatca ccaatcagtt cagacctggt tgattttgta ctttggaaact gtaccttgga
841 tggttttgtt tattaaga gaaacctgaa gt

```

TBP cDNA sequence (SEQ ID NO: 6)

NM_003194 Homo sapiens TATA box binding protein (TBP), mRNA

```

1 ggttcgctgt ggcggggcgc tgggcccgcg gctgtttaac ttcgcttcgg ctggcccata
61 gtgatctttg cagtgaacca gcagcatcac tgtttcttgg cgtgtgaaga taacccaagg
121 aattgaggaa gttgctgaga agagtgtgct ggagatgtct taggaaaaaa ttgaatagtg
181 agacgaagtc cagcgcaagg gtttctggtt tgccaagaag aaagtgaaca tcatggatca
241 gaacaacagc ctgccacctt acgctcaggg cttggcctcc cctcagggtg ccatgactcc
301 cggaatccct atcttttagt caatgatgcc ttatggcact ggactgacct cacagcctat
361 tcagaacacc aatagtctgt ctattttgga agagcaacaa aggcagcagc agcaacaaca
421 acagcagcag cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca
481 gcagcagcag cagcagcagc agcagcagca gcaacaggca gtggcagctg cagccgttca
541 gcagtcaacg tcccagcagg caacacaggg aacctcaggc caggcaccac agctcttcca
601 ctcacagact ctcaacaact cacccttgcc gggcaccact ccaactgtat cctcccccat
661 gactcccatg acccccatca ctectgccac gccagcttcg gagagtctct ggattgtacc
721 gcagctgcaa aatattgtat ccacagtga tcttggttgt aaacttgacc taaagaccat
781 tgcaacttcgt gccgaaacg ccgaatataa tcccaagcgg ttgctgctgg taatcatgag
841 gataagagag ccacgaacca cggcactgat ttctagttct gggaaaatgg tgtgcacagg
901 agccaagagt gaagaacagt ccagactggc agcaagaaaa tatgctagag ttgtacagaa
961 gttgggtttt ccagctaaagt tcttggaact caagattcag aatatggttg ggagctgtga
1021 tgtgaagttt cctataaggt tagaaggcct tgtgctcacc caccaacaat ttagtagtta
1081 tgagccagag ttatttctcg gtttaactca cagaatgatc aaaccagaa ttgttctcct
1141 tatttttgtt tctggaaaag ttgtattaac aggtgctaaa gtcagagcag aaatttatga
1201 agcatttgaa aacatctacc ctattctaaa gggattcagg aagacgacgt aatggctctc
1261 atgtaccctt gcctcccca ccccttctt tttttttttt taacaaaatc agtttgtttt
1321 ggtaccttta aatggtggtg ttgtgagaag atggatgttg agttgcaggg tgtggcacca
1381 ggtgatgcc tctgttaagt gccaccggcg ggatgcccgg aaggggcatt atttgtgcac
1441 tgagaacacc gcgcagcgtg actgtgagtt gctcataccg tgctgctatc tgggcagcgc
1501 tgccatttta tttatatgta gattttaaac actgctgttg acaagtgttg ttgagggaga
1561 aaactttaag tgtaaaagcc acctctataa ttgattggac tttttaattt taatgttttt
1621 cccattgaac cacagttttt atatttctac cagaaaagta aaaatctttt taaaagtgt
1681 tgtttttcta atttataact cctagggttt atttctgtgc cagacacatt ccacctctcc
1741 agtattgcag gacagaatat atgtgttaat gaaaatgaat ggctgtacat atttttttct

```

1801 ttcttcagag tactctgtac aataaatgca gtttataaaa gtgttaaaaa aaaaaaaaaa
1861 aaaaaaa

PDCD2 cDNA sequence 1 (SEQ ID NO: 7)

NM_002598 Homo sapiens programmed cell death 2 (PDCD2), transcript variant 1, mRNA

```

1  tcttgccttc cggcccgcg cccgatttcc gccttccgac ccagctgtgg gctgcgcccc
61  acgccagccc gcgcccgcga tggctgccgc cggggccagg cctgtggagc tgggcttcgc
121 cgagtcggcg cggcggtggc gactgcgcag cgagcagttc cccagcaagg tggcgggcg
181 gccgcatgg ctgggcgcgg cgggctgcc ggggccccag gccctggcct gcgagctgtg
241 cggccgcccg ctctccttcc tgctgcaggt gtatgcgccg ctgcctggcc gcccggaacg
301 cttccaccgc tgcattctcc tcttctgctg ccgcgagcag ccgtgctgtg ccggcctgcg
361 agtttttagg aatcaactac ccaggaaaaa cgatttttac tcatatgagc caccttctga
421 gaatcctccc ccagaaacag gagaatcagt gtgtctccag cttaagtctg gtgctcatct
481 ctgcagggtt tgtggctgtt taggccccaa aacgtgctcc agatgccaca aagcatatta
541 ctgcagcaag gagcatcaga ccctagactg gagattggga cataagcagg cttgtgcaca
601 accagatcat ctggaccata taattccaga ccacaacttc ctttttccag aatttgaaat
661 tgtaatagaa acagaagatg agattatgcc tgaggttgtg gaaaaggaag attactcaga
721 gattataggg agcatgggtg aagcacttga ggaagaactg gattccatgg caaaacatga
781 atccagggaa gataaaattt ttcagaagtt taaaactcag atagcccttg aaccagaaca
841 gattcttaga tatggcagag gtattgcccc catctggatt tctggtgaaa atattcctca
901 agaaaaggat attccagatt gccctgtgg tgccaagaga atattggaat tccaggatcat
961 gcctcagctc ctaaaactacc tgaaggctga cagactgggc aagagcattg actggggcat
1021 cctggctgtc ttcacctgtg ctgagagctg cagcttgggt actggctata cagaagaatt
1081 tgtgtggaag caggatgtaa cagatacacc gtaaaggcat cttaaagcct tgaaaaatgt
1141 taataatctt ttataccttg caattccatt tctgggattt tatcctaagg aaatacttat
1201 accaaaaata gaggtgcaga gatgttgaca gattgcttac acagtgtcta cttattagtg
1261 aaacaaaagt gtcagtgac agggaattaa ataaattttg gtacatccac a

```

PDCD2 cDNA sequence 2 (SEQ ID NO: 8)

NM_144781 Homo sapiens programmed cell death 2 (PDCD2), transcript variant 2, mRNA

```

1  tcttgccttc cggcccgcg cccgatttcc gccttccgac ccagctgtgg gctgcgcccc
61  acgccagccc gcgcccgcga tggctgccgc cggggccagg cctgtggagc tgggcttcgc
121 cgagtcggcg cggcggtggc gactgcgcag cgagcagttc cccagcaagg tggcgggcg
181 gccgcatgg ctgggcgcgg cgggctgcc ggggccccag gccctggcct gcgagctgtg
241 cggccgcccg ctctccttcc tgctgcaggt gtatgcgccg ctgcctggcc gcccggaacg
301 cttccaccgc tgcattctcc tcttctgctg ccgcgagcag ccgtgctgtg ccggcctgcg
361 agtttttagg aatcaactac ccaggaaaaa cgatttttac tcatatgagc caccttctga
421 gaatcctccc ccagaaacag gagaatcagt gtgtctccag cttaagtctg gtgctcatct
481 ctgcagggtt tgtggctgtt taggccccaa aacgtgctcc agatgccaca aagcatatta
541 ctgcagcaag gagcatcaga ccctagactg gagattggga cataagcagg cttgtgcaca
601 accagatcat ctggaccata taattccaga ccacaacttc ctttttccag aatttgaaat
661 tgtaatagaa acagaagatg agattatgcc tgaggttgtg gaaaaggaag attactcaga
721 gattataggg agcatgggtg agcagtttca ggacttcatt cattaagtgg ttaaacataa
781 tacttgggaag aaagggtccc atgtgcctag aagagaggta ctgagaggaa gactcacttt
841 ggaggtctga gcatacaatt ttcagatatt gcctcaggta aaaatatact tcttggactt
901 tgttttctga cacataagag gtgtgttctg ctccctgtaa agacaagggt gggtatccag
961 atgggtcccat gagtagggct gcacaagatg ctggaggctt ggtaagttcc tctgggtcgc
1021 agatcggttt ctgggtcgg gatagtgtga gtgcctagca cagtgtcggg cacgcagaag
1081 ggcccttaa aagtttctct ttcattctggc cagttttaga tacacaattt tgtcagttta
1141 cttacagtgc atactcttgg gtagtacttg tgctgaccaa gtatcttaga ggcttatttt
1201 attatagtag ccaacattta tccagactt acccttatata aagggtctgt tgtgcatgag
1261 ctcatataaaa tcgtgacagc agaccaatga gtgagaaact gccccatttt gaaggtgagg
1321 aaattgaggt tctgggtata actttctttt gtcacataat attaaatttt acaatttgag
1381 ccttgagcca tacacaaaac caccacaaaa ttagatttat agactcaaaa tgaaaacatc

```

```

1441 agcttacttg tttgtagtgc ataccagtcac tacattccaa aacatgtttt gagtcttact
1501 ctgtgcctga ccttgtgctt gataacaggg atataatggg aagcaacact ccagtgggtca
1561 gatgctcaca gtcttatgga ggagcccaaa taatatcttg ggaagttaaa gtccatataa
1621 tgactgataa gagtacaata cagggtgccat gggaacacgt gacatcactg aagactgcct
1681 ggaagggggc gcgcgtgtgt tcatgcctat acgataaaca tgatacataa tgaaaatgct
1741 tatctttagg agaaaggaga gcctagagta gcaggatcaa ggatgaaagc tggacttcaa
1801 atatgccttg ttagtgtaaa tgtgactgtg gaactgtatg agtattttta gattatggag
1861 taaagtaagt tttaaaaagc agtccctaata catcaaaagt aaaaaactct tgatgtagtc
1921 atataaccac actaagaact cttccagggtg acttcaaaac ataggacagt acatctctag
1981 tagaatatgc cctgagaatg aaaagaatgt aacagtgtta gtattttgaa taaacatggt
2041 attactaaaa aaaaaaaaaa aaaaaa

```

ELP3 cDNA sequence (SEQ ID NO: 9)

NM_018091 Homo sapiens elongation protein 3 homolog (S. cerevisiae) (ELP3), mRNA

```

1 gcagaaatga ggcagaagcg gaaaggagat ctcagccctg ctgagctgat gatgctgact
61 ataggagatg ttattaaaca actgattgaa gccacagagc aggggaaaga catcgatcta
121 aataaggtga aaaccaagac agctgccaaa tatggccttt ctgcccagcc ccgcctgggtg
181 gatatacttg ctgccgtccc tccctcagtat cgcaaggctc tgatgcccaa gttaaaggcg
241 aaacccatca gaactgctag tgggattgct gtcgtggctg tgatgtgcaa accccacaga
301 tgtccacaca tcagttttac aggaaatata tgtgtatact gccctgggtg acctgattct
361 gattttgagt attccaccca gtcttacact ggctatgagc caacctccat gagagctatc
421 cgtgccagat atgacccttt cctacagaca agacaccgaa tagaacagtt aaaacaactt
481 ggtcatagtg tggataaagt ggagtttatt gtgatgggtg gaacgtttat ggcccttcca
541 gaagaatata gagattattt tattcgaaat ttacatgatg ccttatcagg acatacttcc
601 aacaatattt acgaggcagt caagtattct gagagaagcc tcacaaagtg tattggaatt
661 actattgaaa ccagaccaga ttactgcatg aagcgacatt taagtgcacat gttgacctat
721 ggctgcacaa ggctggagat tgggggtgcag agtggtttatg aagatgtggc tagagacacc
781 aacagggggc acactgtgaa ggcagtgtgt gagtcatttc acctggccaa agattccggg
841 tttaaagtgg tggcccatat gatgcctgac ctgccaaacg tgggactaga aagagacatt
901 gaacagttca cagagttttt tgagaaccct gcttttcgtc ccgatgggct gaaactctat
961 cctaccctgg tgattcgtgg gaccgggctt tatgagcttt ggaaatcagg aagatataag
1021 agttactctc ctagtgcctt ggttgaattg gtggctcgga tcctagccct cgtgcctcca
1081 tggactcgag gttaccgagt acagagggat attccaatgc ctttagttag ctcaggagta
1141 gagcatggta acctgagaga gctggcactt gcaagaatga aagacctcg aatacagtgt
1201 cgagatgtga gaaccagaga agttggaatc caagaaattc atcacaagat acggccatac
1261 caggttgaat tggtaaggag agattatggt gcaaatgggt gctgggaaac attcttgtca
1321 tacgaagacc catagcaaga cattttgatt ggccctctac gattacgcaa gtgttcagaa
1381 gaaactttcc gtttcgaatt ggggtggagt gtctccatag tacgagagct gcatgtgtat
1441 gggagtgtgg tcctgtgtgag cagccgggat cctactaaat ttcagcatca gggatttggc
1501 atgctgctga tggaggaagc agaaagaata gctagagaag aacatgggtc tgggaaatc
1561 gctgtgatat caggggtcgg caccaggaat tattatagaa agatcggcta cagattacaa
1621 ggcccgtaga tggtagaatg gctgaaataa tggccacacc agtccactct tctgcagtat
1681 cctccctggc agaacacgga gaatcaggat ttcttaataa ctcaacagag aggcctgagca
1741 gagcaaatgg ggggcttcac cctcatcccg cagctgcaga gactggaaac tgccttcaag
1801 gccacggctg gtcattctgt gaccacaccc cagatccgcc ctctcctgcy tgcaccccaa
1861 aaaatcactt gcgtttttga ggcttaaatc atctatccag tttctacatt ttgcatgagg
1921 cctgcagggt gcctattttg actcagacgg tgaaaaaagc aaattaactc atttggacac
1981 cataactcat gcaataaaac tgattgtcat tcgaggagca aacttaagag tagtttattt
2041 atataccctg gggacagaaa gtcagggttga aacaggaaaa ccaccagact ctaatctcag
2101 ccctttaacg acatacgcat tggagcgcaa gttaggaaaa tgagcttttg tttcatgga
2161 aatcattctg attacagtgc tgatgtttag aaataaatag cagtgtgact gggaaagagg
2221 aattgcagtt ctgggggtgt gagcctggca gcagccagcc agcagcctct ccaggcggg
2281 agtctacatc cggagacggc gatgacaaag agcttcattc cacattcttt gttatctcta
2341 ctccccacc tcttggaac tacagagcag tgtgggcagc cccaagtgtg gtccccagag
2401 agcgtttggc tttcctgtct gtctatcctg agcgggtgga gtctcaggtt gtgtgccctt
2461 aaatcaagat ttgcttccac agaagccatt acttgcaatt tttttttttt tttctgagaa
2521 agtctcgtg tgctaccag gctggagtgc agtggcgcaa tctcactgca tctccgctt
2581 cccgggttca agcagattct cgcctcagc gctgggatta caggcaccgg gtcaggctgg
2641 ccgctgctaa tttttgtatt tttagtagag atgggggttt caccatattg gtcaggctgg
2701 tctcgaactc ctgacctcag gtgatcaacc caccttggcc tccctaaatg ccgggattac

```

```

2761 aggcattgagc caccgctccc agcctttgat tttttaaggt ggatttttggg tgttataaat
2821 ggagaaaggt aagagttcaa gttcaacccg tgtgtgaaaag caaaacaatg gaaaacagga
2881 ttggcttctt caaaggctcc tcttgttaga ctgcctcttt gaaatttcga ggtaattctac
2941 tttggagact ctgcctggag agggtcagtt cctaagttaa aagcatcgct taaccttggc
3001 tcctgtggca ttttacaagg gtttaaaagg attgattcct ctgaaagggc ctgaaaataa
3061 aaagtcttta acatataaaa aaaaaaaaaa aaaaa

```

CHDC1 cDNA sequence (SEQ ID NO: 10)

NM_015116 Homo sapiens calponin homology (CH) domain containing 1 (CHDC1), mRNA

```

1  ccgcagtcct tagcttcccg gggacaggaa accttcaaga ccgagctgcc acggccgcct
61  ccccgccgcg ccccattct acgcgcctgc ccacaccctc ctcccctcct tccagcgcct
121  ttcggtggag cactgcggca ctcagcccg gctgccgttt tcccctcgcg gggaacgctg
181  tgaccccccc gcaggagcgg cggggcgggg tggggggggc cgggagaaga tggcgacgcc
241  gggaagcgaa ccccaacctt tcgtcccggc cctttcggtg gctactctgc acccacttca
301  tcatccccac caccaccacc accaccatca gcaccacgga ggaaccggcg ccccgcgcg
361  ggcgggtggg ggcggcggtg gcagcggggg cttcaacctg cccttgaacc ggggtctgga
421  gcgcgcgctt gaggagcgcg ccaactccgg ggggctgaac ctgagcgcca ggaaattgaa
481  ggaatttccc cgtaccgcag ccccggggca tgacctctcg gacacggcgc aggcagactt
541  atctaaaaac agactggttg aagttccaat ggaattgtgc cattttgtat cactggaat
601  tcttaactctg tatcacaact gtatcagagt cattcctgag gccatcgtaa atctgcagat
661  gctgacttac ctgaacttga gtcgaaatca gctgtccgcc ctgctgcctt gcctgtgtgg
721  tctgcctctc aaagtcttaa tcgcaagtaa caacaaactt ggatcattac cagaagagat
781  aggtcagctc aaacagttaa tggagctgga tgtcagctgc aacgagatca cagcgttgcc
841  ccagcagata ggtcagttga aatctctacg agaactgaat gtcagaagaa attaccttaa
901  agttttacca caagaactag tagatcttcc cttggtaaag tttgactttt cctgcaacaa
961  agtgcctgtg attccaattt gttttagaga gatgaagcag ctgcaagtgt tactacttga
1021  gaataaccct ctgcagtctc ctccagcaca gatttgcaca aagggcacaag ttcatatatt
1081  taagtattctg agcatacaag catgccagat taagacagct gactcccttt atctccacac
1141  catggagagg ccacatttac accagcacgt ggaagatggc aagaaggatt ctgattcggg
1201  agttggaagt gataatggag ataagcgatt atctgccacc gagccttctg acgaagacac
1261  tgttagcctc aatgtgcca tgtaaacatt catggaagaa gaacagatca tcaaggagga
1321  ctgtgtccat cgccttagcc ccgttaaagg ggaatttcat cagggaatttc aaccggagcc
1381  ttcccttttg ggtgacagca ccaactcagg agaagaaaga gaccagttta ctgatagagc
1441  agatggtctc cattcggaat ttatgaacta taaggcaagg gcagaagact gtgaagagct
1501  gttacggata gaagaggatg tgcactggca aactgagggc ataataagtt catccaaaga
1561  tcaggacatg gatatagcaa tgatcgagca gctgagagaa gcagtagatt tgcgtcaaga
1621  tcccaatgga ttaagcacag atattacaga gagaagtgtt taaacctat atcttatggg
1681  atcagcagaa gccttagaat tacaagattc tgcactgaat ggtcaaatac agctggagac
1741  atctccggtg tgtgaggtgc aaagtgatct aacattacag agtaacggga gccagtattc
1801  tccaaatgag attagagaga actcccctgc agtctctcct accacaaaca gcacagctcc
1861  atttggcctg aagcctcgat cagtgtttct aagacctcag agaaatttgg aatctataga
1921  cccgcagttt acaatccgga ggaaaatgga gcagatgaga gaagagaaaag agctggtgga
1981  acaacttcgt gagagcattg agatgagatt gaaggctcag ctacacgaag acctgggggc
2041  agccctcatg gatggtgtcg tcctctgcca tctggtcaac cacatccgcc caccgtcggt
2101  tgcaagcatc catgtcccat caccagcggg tcccaaactt agcatggcca aatgcagaag
2161  aaatgtggaa aacttttttg aagcgtgccg aaaattagga gtaccagagg ctgacctctg
2221  ctctccgtgt gacatcctgc agttggattt tcgtcacatt cgaaagactg ttgacactct
2281  gctggcactc ggggagaaag cccaccacc aacttctgcc ctccgctcca gggaccttat
2341  aggtctctgt cttgtccata ttctctttat agtgctggtc tatatcactt accactggaa
2401  tgctctgttc gcataacgtc tgcacgtgca tccaaacgct gtgctctgtc gccctcaacc
2461  tttgcagggt ccttccctacc tttgagcctt tgccctgcaa acttccatcc ctgtcatgtc
2521  ttcagttatc tctcgagttt tgaagctgaa cagtagcaaa tcagattttc cagaagcaca
2581  aactttgtag aatacagttt agtataattc ctctcactta ctgaaataca acgacgcaga
2641  ctgcaaagtg tatgcacacc gcatgcttcc tcatccacat agtgccagca gcagtgccac
2701  gcagttcctc ctctccctcc cggtagagtc ctgccctggg cagaggggag gagaattcca
2761  ggacaagagt gtcaaggaca gggatttagc atatggaagt ctttcccttg ggtcagtatt
2821  gaactagaat tctaattcgg gactgggcaa ttgagctgta taggggccac cttgcaggga
2881  ggacagaaaa ctaacatttt ggcccaactt gatctataca aaactttaat aataccacta
2941  ctgaccaagt tggacgtgta cacgtactca cactgccttg atggccattc gattggattc

```

```

3001 ctcccaaatt tcttaaaaaa ggagccgcga agggcgctgg gcagtgtggc cgccaacttc
3061 caccgccgca agccctctg tccatgacag aagggcgctc caggggaagga agtgtcgttg
3121 ctgttagagc ctcacgtgga ggagtcactt aaacaccagt tttttactgc ttaattcctt
3181 gttaggtctt ctcttgaggc tcttagaaaa gcgttttcca gagagatttc tatttttgaa
3241 caatggaacg gatcactgct tttttgccac atcacatagt aactgccggt ccagaatgtg
3301 acggattcga ctctattcat tttcaataaa agccatgagc cgtggaacat tcttggctct
3361 ggtgcttggg ttatgatggc aggagtcagg aagaagatta ctttcattct agaagaatgt
3421 agtttctcta attatttgaa atgttcattt agcctttgat tttcactgat attaactagc
3481 aaactgcttt aagtcagctc aaaggattat atagtaacta tatctgcatt tggagcaatg
3541 tgatcagttt gcatttaaaa ggaaaaaaa gaattttatc ttagccagaa tgtccctgga
3601 ttcagggtg tctttgtata atatgagagg gccttgttcc aaggtcaagg cagcctcctt
3661 attttacatg ctgtttgcc aatcttgttt cttagcttgg ggagatgatg gacttagctt
3721 cctcaagata aatttctagt ttattaagat gcaaacagct ctcatagatg gctactacga
3781 agaaaaatctt attttctga acattttcat gaatccaggg gacttgaaaa tatggaagac
3841 ccacatagtt agaagaatat atttataaag attccttgct gctaagtcag atcagatttg
3901 ctaacaggaa gcattcttta catgacagta tcttgagtta tgtgagtttt ttttctcctt
3961 gtaattgtgt tgattgtgta aatgcagggt atgtggaagt tatctaatta acctcagttg
4021 tatatgaata acccacagat gtactgaatt acttttggtg ctatcttgta ctcttcaatc
4081 tgtaacacaa taaaatccct ttgtacgatg tctaattgagc accctgagcc ataaattgct
4141 taataaacac attttggtg att

```

ERG cDNA sequence 1 (SEQ ID NO: 11)

NM_182918 Homo sapiens v-ets erythroblastosis virus E26 oncogene like (avian), (ERG), transcript variant 1, mRNA

```

1 aatctcatcc gctctaaaca acctcatcaa aactactttc tggtcagaga gaagcaataa
61 ttattattaa catttattaa cgatcaataa acttgattgc attatggcca gcactattaa
121 ggaagcctta tcagttgtga gtgaggacca gtcgttggtt gagtgtgcct acggaacgcc
181 acacctggct aagacagaga tgaccgcgtc ctctccagc gactatggac agacttccaa
241 gatgagccca cgcgtccctc agcaggattg gctgtctcaa cccccagcca gggtcaccat
301 caaaatggaa tgtaacctta gccaggtgaa tggctcaagg aactctcctg atgaatgcag
361 tgtggccaaa ggcggaaga tgggtggcag cccagacacc gttgggatga actacggcag
421 ctacatggag gagaagcaca tgccaccccc aaacatgacc acgaacgagc gcagagttat
481 cgtgccagca gatcctacgc tatggagtac agaccatgtg cggcagtggtc tggagtgggc
541 ggtgaaagaa tatggccttc cagacgtcaa catcttgtaa ttccagaaca tcgatgggaa
601 ggaactgtgc aagatgacca aggacgactt ccagaggctc acccccagct acaatgccga
661 catccttctc tcacatctcc actacctcag agagactcct ctccacatt tgacttcaga
721 tgatgttgat aaagccttac aaaactctcc acggttaatg catgctagaa acacaggggg
781 tgcagctttt attttccaa atacctcagt atatcctgaa gctacgcaa gaattacaac
841 taggccagat ttaccatatg agccccccag gagatcagcc tggaccggtc acggccaccc
901 cagcggccag tcgaaagctg ctcaaccatc tccttccaca gtgccccaaa ctgaagacca
961 gcgtcctcag ttagatcctt atcagattct tggaccaaca agtagccgcc ttgcaaattc
1021 aggcagtggc cagatccagc tttggcagtt cctcctggag ctctgtcgg acagctccaa
1081 ctccagctgc atcacctggg aaggcaccaa cggggagttc aagatgacgg atccgacga
1141 ggtggccgg cgctggggag agcggaagag caaacccaac atgaactacg ataagctcag
1201 ccgcgccctc cgttactact atgacaagaa catcatgacc aagggtccatg ggaagcgcta
1261 cgctacaag ttcgacttcc acgggatcgc ccaggccctc cagccccacc cccggagtc
1321 atctctgtac aagtaccctc cagacctccc gtacatgggc tcctatcacg cccaccaca
1381 gaagatgaac tttgtggcgc cccacctcc agccctcccc gtgacatctt ccagttttt
1441 tgctgcccc aaccataact ggaattcacc aactgggggt atatacccca acactaggct
1501 cccaccagc catatgcctt ctcatctggg cacttactac taaagacctg gcggaggctt
1561 ttcccatcag cgtgcattca ccagcccatc gccacaaact ctatcggaga acatgaatca
1621 aaagtgcctc aagaggaatg aaaaaagctt tactggggct ggggaaggaa gccggggaag
1681 agatccaaag actcttggga gggagtact gaagtcttac tacagaaatg aggaggatgc
1741 taaaaatgtc acgaatatgg acatatcatc tgtggactga ccttgtaaaa gacagtgtat
1801 gtagaagcat gaagtcttaa ggacaaagt ccaaagaaag tgggtcttaag aatgtataa
1861 acttttagagt agagtttgga atcccactaa tgcaaactgg gatgaaacta aagcaataga
1921 aacaacacag ttttgacctt acataccgtt tataatgcca ttttaaggaa aactacgtgt
1981 atttaaaaat agaaacatat caaaaaaaaa aaaaaa

```

ERG cDNA sequence 2 (SEQ ID NO: 12)

NM_004449 Homo sapiens v-ets erythroblastosis virus E26 oncogene like (avian), (ERG), transcript variant 2, mRNA

```

1 atgattcaga ctgtcccga cccagcagct catatcaagg aagccttatt agttgtgagt
61 gaggaccagt cgttgtttga gtgtgcctac ggaacgccac acctggctaa gacagagatg
121 accgcgtcct cctccagcga ctatggacag acttccaaga tgagcccacg cgtccctcag
181 caggattggc tgtctcaacc cccagccagg gtcacccatca aaatggaatg taaccctagc
241 caggtgaatg gctcaaggaa ctctcctgat gaatgcagtg tggccaaagg cgggaagatg
301 gtgggcagcc cagacaccgt tgggatgaac tacggcagct acatggagga gaagcacatg
361 ccaccccaaa acatgaccac gaacgagcgc agagttatcg tgccagcaga tcctacgcta
421 tggagtacag accatgtgcg gcagtggctg gagtgggcgg tgaaagaata tggccttcca
481 gacgtcaaca tcttgttatt ccagaacatc gatgggaagg aactgtgcaa gatgaccaag
541 gacgacttcc agaggctcac cccagctac aacgccgaca tccttctctc acatctccac
601 tacctcagag agactcctct tccacatttg acttcagatg atgttgataa agccttacia
661 aactctccac ggttaatgca tgc tagaaac acagatttac catatgagcc cccaggaga
721 tcagcctgga ccggtcacgg ccacccacg cccagtcga aagctgctca accatctcct
781 tccacagtgc ccaaaactga agaccagcgt cctcagttag atccttatca gattcttgga
841 ccaacaagta gccgccttgc aaatccaggc agtggccaga tccagctttg gcagttcctc
901 ctggagctcc tgtcggacag ctccaactcc agctgcatca cctgggaagg caccacggg
961 gagtcaaga tgacggatcc cgacgaggtg gcccgcgct ggggagagcg gaagacaaa
1021 ccaacatga actacgataa gctcagccgc gccctccgtt actactatga caagaacatc
1081 atgaccaagg tccatgggaa gcgctacgcc tacaagttcg acttccacgg gatcgccag
1141 gccctccagc cccaccccc ggagtcactc ctgtacaagt acccctcaga cctcccgta
1201 atgggctcct atcacgcca cccacagaag atgaactttg tggcgcccca ccctccagcc
1261 ctccccgtga catcttccag tttttttgct gcccacaaacc catactggaa ttcaccaact
1321 ggggtatat accccaacac taggctcccc accagccata tgccttctca tctgggcact
1381 tactactaa

```

[0270] Following are amino acid sequences for *PSMB1* (SEQ ID NO: 13), *TBP* (SEQ ID NO: 14), *PDCD2* (SEQ ID NO: 15 and 16), *ELP3* (SEQ ID NO: 17), *CHDC1* (SEQ ID NO: 18), and *ERG* (SEQ ID NO: 19 and 20).

PSMB1 amino acid sequence (SEQ ID NO: 13)

NP_002784 Homo sapiens proteasome (prosome, macropain) subunit, beta type, 1 (PSMB1), protein

MLSSTAMYSAPGRDLGMEPHRAAGPLQLRFSPYVFNGGTILAIAGEDFAIVASDTRLSE
GFSIHTRDSPKCYKLTDKTVIGCSGFHGDCLTLTKIIEARLKMYKHSNNKAMTTGAIAA
MLSTILYSRRFFPYVYNIIGGLDEEGKGAVYSFDPVGSYQRDSFKAGGSASAMLPQLL
DNQVGFKNMQNVEHVPLSLDRAMRLVKDVFISAAERDVYTGDALRICIVTKEGIREET
VSLRKD

TBP amino acid sequence (SEQ ID NO: 14)

NP_003185 Homo sapiens TATA box binding protein (TBP), protein

MDQNNSLPPYAQGLASPQGAMTPGIPIFSPMMPYGTGLTPQPIQNTNSLSILEEQQRQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQAVAAAQVQSTSQQATQGT
SGQAPQLFHSQTLTAPLPGTTPLYPSPMTPMTPITPATPASESSGIVPQLQNIVSTVNLG

CKLDLKTIALRARNAEYNPKRFAAVIMRIREPRTTALIFSSGKMVCTGAKSEEQSRLAAR
KYARVVQKLGFPKFLDFKIQNMVGSCDVKFPIRLEGLVLTHQQFSSYEPELFPGLIYR
MIKPRIVLLIFVSGKVVLTGAKVRAEIYEAFENIYPILKGFRKTT

PDCD2 amino acid sequence 1 (SEQ ID NO: 15)

NP_002589 Homo sapiens programmed cell death 2 (PDCD2), isoform 1, protein

MAAAGARPVELGFAESAPAWRLRSEQFPSKVGGPRAWLGAAGLPGPQALACELCGRP
LSFLLQVYAPLPGRPDAFHRCIFLCCREQPCCAGLRVFRNQLPRKNDFYSEPPSENPPP
ETGESVCLQLKSGAHLRCRVCGCLGPKTCSRCHKAYYCSKEHQTLDWRLGHKQACAQP
DHLDHIIIPDHNFLFPEFEIVIETEDIMPEVVEKEDYSEIIGSMGEALEEELD SMAKHESRE
DKIFQKFKTQIALEPEQILRYGRGIAPIWISGENIPQEKDIPDCPCGAKRILEFQVMPQLLN
YLKADRLGKSIDWGILAVFTCAESCSLGTGYTEEFVWKQDVTDTTP

PDCD2 amino acid sequence 2 (SEQ ID NO: 16)

NP_659005 Homo sapiens programmed cell death 2 (PDCD2), isoform 2, protein

MAAAGARPVELGFAESAPAWRLRSEQFPSKVGGPRAWLGAAGLPGPQALACELCGRP
LSFLLQVYAPLPGRPDAFHRCIFLCCREQPCCAGLRVFRNQLPRKNDFYSEPPSENPPP
ETGESVCLQLKSGAHLRCRVCGCLGPKTCSRCHKAYYCSKEHQTLDWRLGHKQACAQP
DHLDHIIIPDHNFLFPEFEIVIETEDIMPEVVEKEDYSEIIGSMGKQFQDFIH

ELP3 amino acid sequence (SEQ ID NO: 17)

NP_060561 Homo sapiens elongation protein 3 homolog (S. cerevisiae) (ELP3), protein

MRQKRKGDLSPAELMMLTIGDVIKQLIEAHEQGKDIDLNKVKTKTAAKYGLSAQPRLV
DIIAAVPPQYRKVLMPKLKAKPIRTASGIAVVAVMCKPHRCPHISFTGNICVYCPGGPDS
DFEYSTQSYTGYEPTSMRAIRARYDPFLQTRHRIEQLKQLGHSVDKVEFIVMGGTFMAL
PEEYRDYFIRNLHDALSGHTSNNIYEAVKYSERSLTKCIGITIETRPDYCMKRHLSDMLT
YGCTRLEIGVQSVYEDVARDTNRGHTVKAVCESFHAKDSGFKVVAHMMPDLPNVGL
ERDIEQFTEFFENPAFRPDGLKLYPTLVIRGTGLYELWKSGRYKSYSPSDLVELVARILA
LVPPWTRVYRVQRDIPMLVSSGVEHGNLRELALARMKDLGIQCRDVRTREVGIQEIHH
KVRPYQVELVRRDYVANGGWETFLSYEDPDQDILIGLLRLRKCEETFRFELGGGVSI
RELHVGSVVPVSSRDPTKFQHQGFGLLMEEAERIAREEHSGSGKIAVISGVGTRNYR
KIGYRLQGPYMKMLK

CHDC1 amino acid sequence (SEQ ID NO: 18)

NP_055931 Homo sapiens calponin homology (CH) domain containing 1 (CHDC1), protein

MATPGSEPQPFVPALSVATLHPLHHPHHHHHHHHQHGGGTGAPGGAGGGGGGGSGGFNL
PLNRGLERALEEAANSGLNLSARKLKEFPRTAAPGHDLSDTVQADLSKNRLVEVPME
LCHFVSLEILNLYHNCIRVIPEAIVNLQMLTYLNLSRNQLSALPACLCGLPLKVLIASNNK
LGSLPEEIGQLKQLMELDVSCNEITALPQQIGQLKSLRELNVRNRYLKVLPQELVDLPLV
KFDFSCNKVLVIPICFREMKQLQVLLLENNPLQSPPAQICTKGKVHIFKYLSIQACQIKTA
DSLYLHTMERPHLHQHVEDGKKDSDSGVGSDNGDKRLSATEPSDEDTVSLNVPMSNM
EEEQIIKEDSCHRLSPVKGEFHQEFQPEPSLLGDSTNSGEERDQFTDRADGLHSEFMNYK
ARAEDCELLRIEEDVHWQTEGISSSKDQDMDIAMIEQLREAVDLLQDPNGLSTDITER
SVLNLYPMGSAEALQLQDSALNGQIQLETSPVCEVQSDTLQSNQSQYSPNEIRENSPAV
SPTTNSTAPFGLKPRSVFLRPQRNLESIDPQFTIRRKMEQMREEKELVEQLRESIEMRLKV
SLHEDLGAALMDGVVLCHLVNHIRPRSVASIHVPSPAVPKLSMAKCRNVENFLEACR
KLGVP EADLCSPCDILQLDFRHIRKTVDTLLALGEKAPPPTSALRSRDLIGFCLVHILFIVL
VYITYHWNALSA

ERG amino acid sequence 1 (SEQ ID NO: 19)

NP_891548 Homo sapiens v-ets erythroblastosis virus E26 oncogene like (avian), (ERG), isoform 1, protein

MASTIKEALSVVSEDQSLFECAYGTPHLAKTEMTASSSSDYGQTSKMSPRVPQQDWLS
QPPARVTIKMECNPSQVNGSRNSPDECSVAKGGKMVGSPDTVGMNYGSYMEEKHMPP
PNMTTNERRVIVPADPTLWSTDHVRQWLEWAVKEYGLPDVNILLFQNIDGKELCKMT
KDDFQRLTPSYNADILLSHLHYLRETPLPHLTSDDVDKALQNSPRLMHARNTGGAAFIF
PNTSVYPEATQRITTRPDLPEPPRRSAWTGHGHPTPQSKAAQSPSTVPKTEDQRPQLD
PYQILGPTSSRLANPGSGQIQLWQFLLELLSDSSNSSCITWEGTNGEFKMTDPDEVARRW
GERKSKPNMNYDKLSRALRYYYDKNIMTKVHGKRYAYKFDHFGIAQALQPHPPESSLY
KYP S DLPYMG SYHAHPQKMNFVAPHPALPVTSSSFFAAPNPYWN SPTGGIYPNTRLPT
SHMPSHLGTYT

ERG amino acid sequence 2 (SEQ ID NO: 20)

NP_004440 Homo sapiens v-ets erythroblastosis virus E26 oncogene like (avian), (ERG), isoform 2, protein

MIQTVDPDAAHIKEALSVVSEDQSLFECAYGTPHLAKTEMTASSSSDYGQTSKMSPRVP
QQDWLSQPPARVTIKMECNPSQVNGSRNSPDECSVAKGGKMVGSPDTVGMNYGSYME
EKHMPPPNMTTNERRVIVPADPTLWSTDHVRQWLEWAVKEYGLPDVNILLFQNIIDGKE
LCKMTKDDDFQRLTPSYNADILLSHLHYLRETPLPHLTSDDDVDKALQNSPRLMHARNTD
LPYEPPRRSAWTGHGHPTPQSKAAQPSSTVPKTEDQRPQLDPYQILGPTSSRLANPGSG
QIQLWQFLLELLSDSSNSSCITWEGTNGEFKMTDPDEVARRWGERKSKPNMNYDKLSR
ALRYYYDKNIMTKVHGKRYAYKFDFHGIAQALQPHPPESLYKYPSDLPYMGSYHAHP
QKMNFVAPHPPALPVTSSSFFAAPNPYWNSPTGGIYPNTRLPTSHMPSHLGTYT

[0271] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the aspects which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0272] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.

What is claimed is:

1. A method for identifying a subject at risk of osteoarthritis, which comprises detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variations are detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-12;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(d) a fragment of a nucleotide sequence of (a), (b), or (c);

whereby the presence of the polymorphic variation is indicative of the subject being at risk of osteoarthritis.

2. The method of claim 1, which further comprises obtaining the nucleic acid sample from the subject.

3. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 170719500 to 170766500 in human genomic DNA.

- 4. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 1 selected from the group consisting of 229, 6310, 11840, 11870, 12064, 13392, 16354, 16559, 16935, 17616, 17737, 18321, 18453, 18811, 20020, 21662, 23197, 23446, 24339, 25504, 27174, 28008, 29294, 29759, 30832, 44512, 44850, 45884, 46345, 48589, 53371, 53911, 53990, 55152, 55667, 58952, 59315, 60029, 61477, 62988, 63090, 64021, 65685, 70220, 70323, 70959, 73436, 82945, 82958, 82961, 82964, 82965, 83006, 83025, 83034, 83074, 83132, 83155, 83172, 83174, 83206, 83216, 83234, 83252, 83260, 83263, 83296, 83319, 83322, 83324, 83357, 83375, 83381, 83389, 83443, 83499, 83545, 83566, 83591, 83619, 83698, 83780, 83784, 83826, 83832, 83852, 86297, 86315, 86420, 86460, 86714, 86718, 86736, 86753, 86766, 88162, 88218, 88246, 88255, 88309, 88310, 88471, 88619, 88904, 89044, 90531, 90534, 90613 and 46252.

5. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 1 selected from the group consisting of 229, 6310, 16559, 18453, 25504, 27174, 30832, 44850, 45884, 48589, 61477, 82961 and 46252.

6. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 27963000 to 27983000 in human genomic DNA.

7. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 2 selected from the group consisting of 211, 473, 1536, 5639, 17186, 17335, 25029, 25111, 28811, 28863, 30809, 40985, 45147, 45282, 46168, 46328, 49077, 51925, 52141, 52168, 60852, 62468, 65572, 79089, 79541, 79790, 90843, 90978, 91052, 91131, 91132, 94439 and 94621.

8. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 2 selected from the group consisting of 40985, 46168, 51925 and 52168.

9. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 44962000 to 45013000 in human genomic DNA.

10. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 3 selected from the group consisting of 243, 10208, 15049, 15111, 15272, 15287, 15326, 15327, 17038, 19391, 21702, 22431, 22881, 27744, 32564, 32698, 33104, 33181, 33256, 33543, 35567, 40085, 40482, 45641, 46059, 48504, 48919, 49693, 49874, 50020, 50616, 50719, 55511, 65533, 70529, 75591, 77266, 80368, 82475, 92462, 92480, 95819 and 96275.

11. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 3 selected from the group consisting of 15111, 45641, 46059, 49693, 49874, 50020, 50719, 70529, 82475, 92462, 92480 and 96275.

12. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 38830000 to 38844000 in human genomic DNA.

13. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 4 selected from the group consisting of 231, 882, 960, 1194, 1530, 1673, 2096, 2285, 5873, 7256, 7988, 8222, 8381, 8814, 8915, 9642, 9902, 10619, 10927, 11032, 14377, 15608, 15928, 16296, 17598, 19272, 20084, 20577, 28051, 29466, 29530, 29987, 30012, 30322, 32216, 32516, 32544, 32746, 33137, 33538, 33798, 33802, 33964, 34132, 34210, 34317, 34499, 34753, 34845, 35335, 36423, 36450, 36481, 38447, 38784, 39387, 39458, 39822, 40305, 40869, 40926, 41010, 41134,

41984, 42172, 42753, 43011, 43176, 43320, 43381, 44142, 44383, 44726, 45087, 45141, 45359, 45421, 45456, 45467, 45486, 45709, 45716, 47626, 49413, 49796, 49962, 50075, 50093, 50571, 50615, 50780, 50851, 51459, 53193, 53702, 53736, 53795, 54109, 54126, 54230, 54894, 55455, 55499, 56522, 56662, 56954, 57267, 58282, 58916, 59544, 59666, 59913, 66846, 67245, 67652, 67955, 67966, 68420, 70226, 70810, 72246, 73330, 73457, 74389, 74638, 74640, 75358, 75952, 76098, 77836, 78449, 78507, 80031, 81695, 82775, 82795, 84611, 84657, 84693, 85020, 85048, 85100, 85325, 85452, 85868, 85936, 85990, 86139, 86497, 87236, 87248, 87533, 87912, 88108, 88494, 89598, 90235, 91287, 91359, 92384, 92410, 92900, 94495, 94512, 97777 and 98333.

14. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 4 selected from the group consisting of 1673, 20577, 33137, 39822, 45716, 49962, 51459, 54894, 55455, 55499, 58282, 68420 and 80031.

15. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in linkage disequilibrium with one or more positions in claim 4, 7, 10 or 13.

16. The method of claim 1, wherein detecting the presence or absence of the one or more polymorphic variations comprises:

hybridizing an oligonucleotide to the nucleic acid sample, wherein the oligonucleotide is complementary to a nucleotide sequence in the nucleic acid and hybridizes to a region adjacent to the polymorphic variation;

extending the oligonucleotide in the presence of one or more nucleotides, yielding extension products; and

detecting the presence or absence of a polymorphic variation in the extension products.

17. The method of claim 1, wherein the subject is a human.

18. The method of claim 17, wherein the subject is a human female.

19. The method of claim 17, wherein the subject is a human male.

20. A method for identifying a polymorphic variation associated with osteoarthritis proximal to an incident polymorphic variation associated with osteoarthritis, which comprises:

identifying a polymorphic variation proximal to the incident polymorphic variation associated with osteoarthritis, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-12;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation;

determining the presence or absence of an association of the proximal polymorphic variant with osteoarthritis.

21. The method of claim 20, wherein the incident polymorphic variation is at one or more positions in claim 4, 7, 10 or 13.

22. The method of claim 20, wherein the proximal polymorphic variation is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the incident polymorphic variation.

23. The method of claim 20, which further comprises determining whether the proximal polymorphic variation is in linkage disequilibrium with the incident polymorphic variation.

24. The method of claim 20, which further comprises identifying a second polymorphic variation proximal to the identified proximal polymorphic variation associated with osteoarthritis and determining if the second proximal polymorphic variation is associated with osteoarthritis.

25. The method of claim 24, wherein the second proximal polymorphic variant is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the proximal polymorphic variation associated with osteoarthritis.

26. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-12;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

(e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);

wherein the nucleotide sequence comprises a polymorphic variation associated with osteoarthritis selected from the group consisting of in SEQ ID NO: 1 a thymine at position 229, a guanine at position 6310, a thymine at position 16559, an adenine at position 18453, an adenine at position 25504, an adenine at position 27174, an adenine at position 30832, a guanine at position 44850, an adenine at position 45884, an adenine at position 48589, a cytosine at position 61477, a cytosine at position 82961 and a thymine at position 46252; in SEQ ID NO: 2 a cytosine at position 40985, a guanine at position 46168, a thymine at position 51925 and a cytosine at position 52168; in SEQ ID NO: 3 a guanine at position 15111, a thymine at position 45641, an adenine at position 46059, a cytosine at position 49693, an adenine at position 49874, an adenine at position 50020, a guanine at position 50719, an adenine at position 70529, an adenine at position 82475, a thymine at position 92462, a thymine at position 92480 and a cytosine at position 96275; and in SEQ ID NO: 4 a guanine at position 1673, a thymine at position 20577, a guanine at position 33137, a guanine at position 39822, an adenine at position 45716, a guanine at position 49962, an adenine at position 51459, a cytosine at position 54894, an adenine at position 55455, an adenine at position 55499, a guanine at position 58282, an adenine at position 68420 and a thymine at position 80031.

27. An oligonucleotide comprising a nucleotide sequence complementary to a portion of the nucleotide sequence of (a), (b), (c), or (d) in claim 26, wherein the 3' end of the oligonucleotide is adjacent to a polymorphic variation associated with osteoarthritis.

28. A microarray comprising an isolated nucleic acid of claim 26 linked to a solid support.

29. An isolated polypeptide encoded by the isolated nucleic acid sequence of claim 26.

30. A method for identifying a candidate therapeutic for treating osteoarthritis, which comprises:

(a) introducing a test molecule to a system which comprises a nucleic acid comprising a nucleotide sequence selected from the group consisting of:

(i) a nucleotide sequence in SEQ ID NO: 1-12;

(ii) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(iv) a fragment of a nucleotide sequence of (a), (b), or (c); or

introducing a test molecule to a system which comprises a protein encoded by a nucleotide sequence of (i), (ii), (iii), or (iv); and

(b) determining the presence or absence of an interaction between the test molecule and the nucleic acid or protein,

whereby the presence of an interaction between the test molecule and the nucleic acid or protein identifies the test molecule as a candidate therapeutic for treating osteoarthritis.

31. The method of claim 30, wherein the system is an animal.

32. The method of claim 30, wherein the system is a cell.

33. The method of claim 30, wherein the nucleotide sequence comprises one or more polymorphic variations associated with osteoarthritis.

34. The method of claim 33, wherein the one or more polymorphic variations associated with osteoarthritis are at one or more positions in claim 4, 7, 10 or 13.

35. A method for treating osteoarthritis in a subject, which comprises contacting one or more cells of a subject in need thereof with a nucleic acid, wherein the nucleic acid comprises a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-12;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(d) a fragment of a nucleotide sequence of (a), (b), or (c); and

(e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);

whereby contacting the one or more cells of the subject with the nucleic acid treats the osteoarthritis in the subject.

36. The method of claim 35, wherein the nucleic acid is RNA or PNA.

37. The method of claim 36, wherein the nucleic acid is duplex RNA.

38. A method for treating osteoarthritis in a subject, which comprises contacting one or more cells of a subject in need thereof with a protein, wherein the protein is encoded by a nucleotide sequence which comprises a polynucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-12;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(d) a fragment of a nucleotide sequence of (a), (b), or (c);

whereby contacting the one or more cells of the subject with the protein treats the osteoarthritis in the subject.

39. A method for treating osteoarthritis in a subject, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variation are detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-12;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

administering an osteoarthritis treatment to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

40. The method of claim 39, wherein the one or more polymorphic variations are detected at one or more positions in claim 4, 7, 10 or 13.

41. The method of claim 39, wherein the treatment is selected from the group consisting of administering a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondrotin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

42. A method for detecting or preventing osteoarthritis in a subject, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-12;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

administering an osteoarthritis prevention or detection procedure to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

43. The method of claim 42, wherein the one or more polymorphic variations are detected at one or more positions in claim 4, 7, 10 or 13.

44. The method of claim 42, wherein the osteoarthritis prevention is selected from the group consisting of administering a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondrotin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

45. A method of targeting information for preventing or treating osteoarthritis to a subject in need thereof, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-12;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-12;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-12;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

directing information for preventing or treating osteoarthritis to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

46. The method of claim 45, wherein the one or more polymorphic variations are detected at one or more positions in claim 4, 7, 10 or 13.

47. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and an antibody that specifically binds to a protein, polypeptide or peptide encoded by a nucleotide sequence identical to or 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-12.

48. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and a RNA, DNA, PNA or ribozyme molecule comprising a nucleotide sequence identical to or 90% or more identical to a portion of a nucleotide sequence in SEQ ID NO: 1-12.

49. The composition of claim 48, wherein the RNA molecule is a short inhibitory RNA molecule.

Abstract of the Disclosure

Provided herein are methods for identifying a risk of osteoarthritis in a subject, reagents and kits for carrying out the methods, methods for identifying candidate therapeutics for treating osteoarthritis, and therapeutic and preventative methods applicable to osteoarthritis. These embodiments are based upon an analysis of polymorphic variations in nucleotide sequences within the human genome.

FIGURE 1A

CHROM 6 – DISCOVERY P-VALUES (female only)

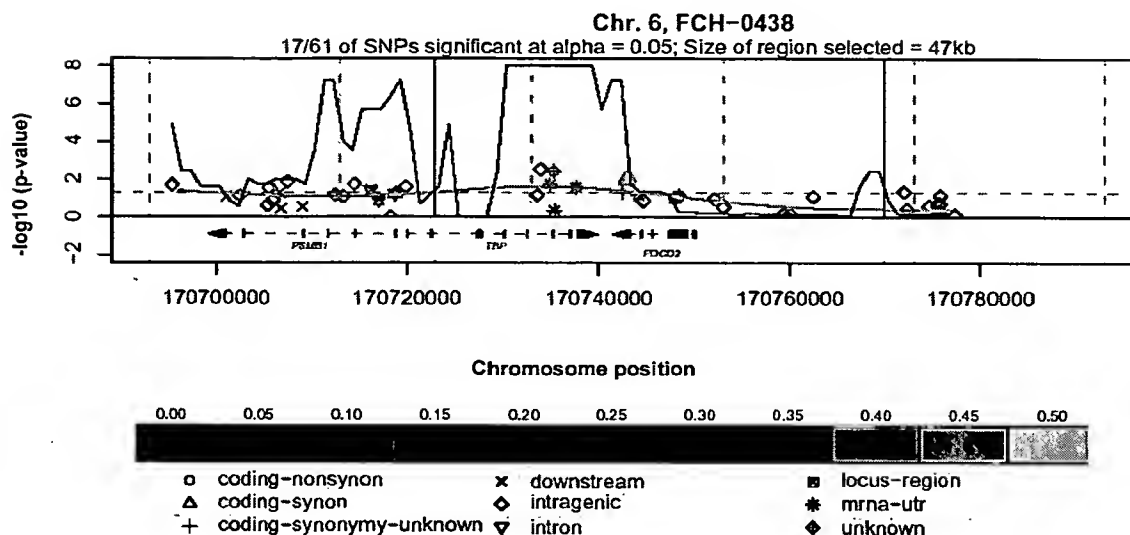


FIGURE 1B

ELP3 – DISCOVERY P-VALUES (female only)

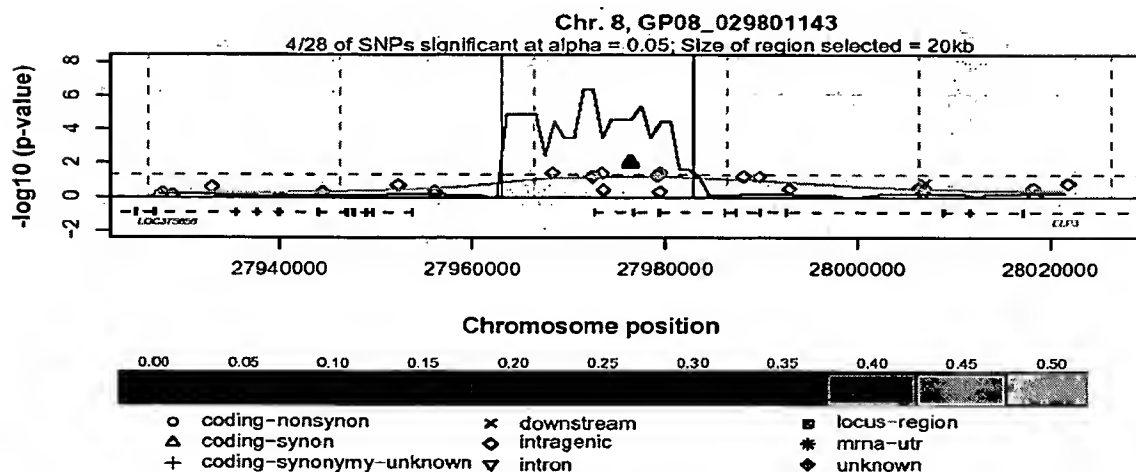


FIGURE 1C

CHDC1 – DISCOVERY P-VALUES (female only)

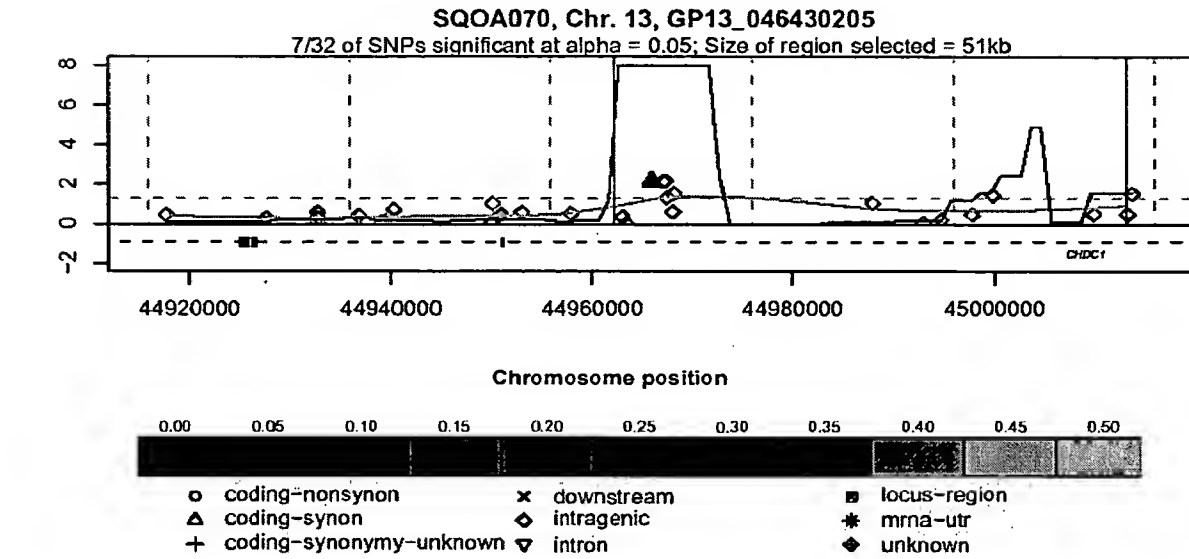
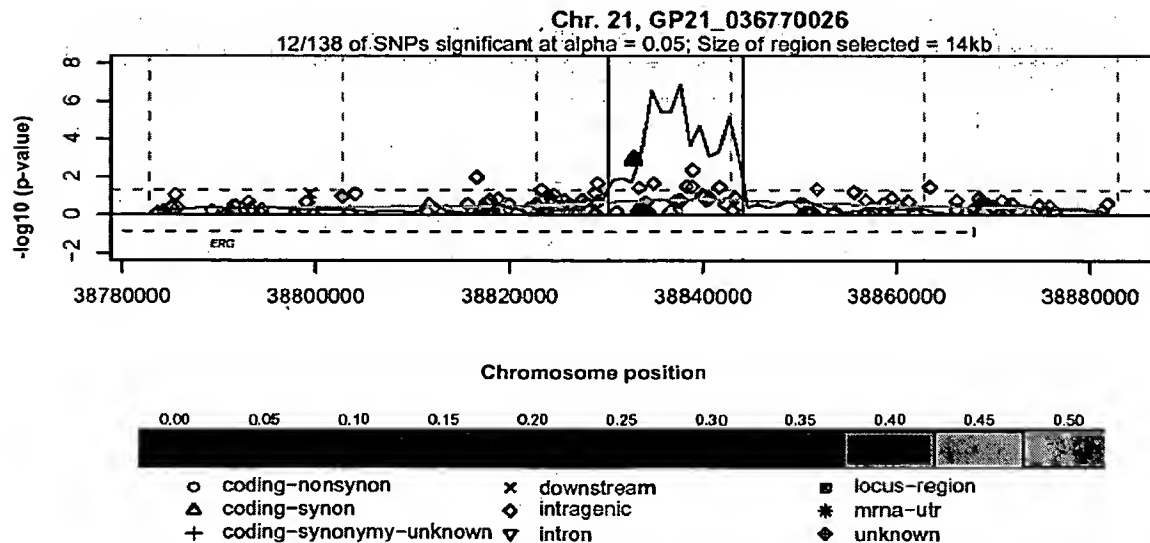


FIGURE 1D

ERG – DISCOVERY P-VALUES (female only)



Application Data Sheet

Application Information

Application Type::	Provisional
Subject Matter::	Utility
Suggested Group Art Unit::	Not Yet Assigned
CD-ROM or CD-R?::	None
Sequence submission?::	None
Computer Readable Form (CRF)?::	No
Title::	METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF
Attorney Docket Number::	524593008700
Request for Early Publication?::	No
Request for Non-Publication?::	No
Total Drawing Sheets?::	2
Small Entity?::	Yes
Petition included?::	No
Secrecy Order in Parent Appl.?::	No

Applicant Information

Applicant Authority Type::	Inventor
Primary Citizenship Country::	US
Status::	Full Capacity
Given Name::	Steven
Family Name::	MAH
City of Residence::	San Diego
State or Province of Residence::	CA
Country of Residence::	US
Street of mailing address::	12820 Via Nieve #74
City of mailing address::	San Diego
State or Province of mailing address::	CA

Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Germany
Status:: Full Capacity
Given Name:: Andreas
Family Name:: BRAUN
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3935 Lago Di Grata Circle
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Germany
Status:: Full Capacity
Given Name:: Stefan
Middle Name:: M.
Family Name:: KAMMERER
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3825 Elijah Court, Unit 334
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: US
Status:: Full Capacity

Given Name:: Matthew
Middle Name:: Roberts
Family Name:: NELSON
City of Residence:: San Marcos
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 1250 Calle Prospero
City of mailing address:: San Marcos
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92069

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Sweden
Status:: Full Capacity
Given Name:: Rikard
Middle Name:: Henry
Family Name:: RENELAND
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 7555 Charmant Drive, #1114
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92122

Applicant Authority Type:: Inventor
Primary Citizenship Country:: United Kingdom
Status:: Full Capacity
Given Name:: Maria
Middle Name:: L.
Family Name:: LANGDOWN
City of Residence:: San Diego

State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3701 Yosemite Street
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92109

Correspondence Information

Correspondence Customer Number:: 25225

Representative Information

Representative Customer Number:: 25225